

# A HYBRID HDMR FOR MIXED MULTISCALE FINITE ELEMENT METHOD WITH APPLICATION FOR FLOWS IN RANDOM POROUS MEDIA \*

LIJIAN JIANG<sup>†</sup>, J. DAVID MOULTON<sup>‡</sup>, AND JIA WEI<sup>§</sup>

**Abstract.** Stochastic modeling has become a popular approach to quantify uncertainty in flows through heterogeneous porous media. The uncertainty in heterogeneous structure properties is often parameterized by a high-dimensional random variable. This leads to a deterministic problem in a high-dimensional parameter space and the numerical computation becomes very challengeable as the dimension of the parameter space increases. To efficiently tackle the high-dimensionality, we propose a hybrid high dimensional model representation (HDMR) technique, through which the high-dimensional stochastic model is decomposed into a moderate-dimensional stochastic model in a most active random space and a few one-dimensional stochastic models. The derived low-dimensional stochastic models are solved by incorporating sparse grid stochastic collocation method into the proposed hybrid HDMR. The porous media properties such as permeability are often heterogeneous. To treat the heterogeneity, we use a mixed multiscale finite element method (MMsFEM) to simulate each of derived stochastic models. To capture the non-local spatial features of the porous media and the important effects of random variables, we can hierarchically incorporate the global information individually from each of random parameters. This significantly enhances the accuracy of the multiscale simulation. The synergy of the hybrid HDMR and the MMsFEM reduces the stochastic model of flows in both stochastic space and physical space, and significantly decreases the computation complexity. We carefully analyze the proposed HDMR technique and the derived stochastic MMsFEM. A few numerical experiments are carried out for two-phase flows in random porous media and support the efficiency and accuracy of the MMsFEM based on the hybrid HDMR.

**Key words.** Hybrid high-dimensional model representation, Sparse grid collocation method, Mixed multiscale finite element method, Approximate global information

**AMS subject classifications.** 65N30, 65N15, 65C20

**1. Introduction.** The modeling of dynamic flow and transport processes in geologic porous media plays a significant role in the management of natural resources such as in oil reservoirs and water aquifers. These porous media are often created by complex geological processes and may contain materials with widely varying abilities to transmit fluids. Thus multiscale phenomena occur in the model. Because of measurement errors and limited knowledge of physical properties, modeling of subsurface flow and transport often uses random fields (e.g., permeability). The model's output can be accurately predicted by efficiently simulating the stochastic multiscale model. The existence of heterogeneity at multiple scales combined with this uncertainty brings significant challenges to the development of efficient algorithms for the simulation of the dynamic processes in random porous media. Therefore, the interest in developing stochastic multiscale methods for stochastic subsurface models has steadily grown in recent years [10, 11, 14, 19, 20, 26, 35].

The existence of uncertainty in random porous media is an important challenge for simulations. One way to describe the uncertainty is to model the random porous

---

\*L. Jiang and J. D. Moulton acknowledge the fund by the Department of Energy at Los Alamos National Laboratory under contracts DE-AC52-06NA25396 and the DOE Office of Science Advanced Computing Research (ASCR) program in Applied Mathematical Sciences.

<sup>†</sup>Applied Mathematics and Plasma Physics, Los Alamos National Laboratory, NM 87545 (ljiang@lanl.gov). Corresponding author.

<sup>‡</sup>Applied Mathematics and Plasma Physics, Los Alamos National Laboratory, NM 87545 (moulton@lanl.gov).

<sup>§</sup>CGGVeritas, Houston, TX 77072 (weijialily@gmail.com)

media as a random field which satisfies certain statistical correlations. This naturally results in describing the flow and transport problem using stochastic partial differential equations (SPDE). Many numerical methods have been developed for solving SPDE in the past decades. The existing stochastic numerical approaches roughly fall into the two categories: (1) non-intrusive schemes and (2) intrusive schemes. In non-intrusive schemes, the existing deterministic solvers are used without any modification to solve a (large) set of deterministic problems, which correspond to a set of samples from the random space. This leads to a set of outputs, which are used to recover the desired statistical quantities. Monte Carlo methods [9] and stochastic collocation methods fall into this category. In intrusive schemes, approximation of the probabilistic dependence is embedded directly in an existing deterministic method. Thus, the resulting numerical model is fundamentally different from the original deterministic model. Typical examples from this class are stochastic Galerkin [27] and perturbation methods [35]. A broad survey of these methods can be found in [22]. Among these methods, stochastic collocation methods have been paid much attention in the research community because they share the fast convergence property of the stochastic finite element methods while having the decoupled nature of Monte Carlo methods. A stochastic collocation method consists of two components: collocation points and interpolation operator. The collocation points are selected based on special rules. The different choice of collocation points leads to the different collocation methods including full-tensor product method [3] and Smolyak sparse grid method [4, 29, 32]. Sparse grid stochastic collocation method is known to have the same asymptotic accuracy as full-tensor product collocation method by using fewer collocation points. However, the number of collocation points required in the sparse grid method will increase drastically for high-dimensional stochastic problems. Hence the method is still limited to the problems in a moderate-dimensional stochastic space.

Due to the small correlation length and large variability of the covariance structure, the uncertainties of random porous media property are usually parameterized by a large number of random variables, e.g., using truncated Fourier type expansion for random fields. Consequently, the model's input is defined in a high-dimensional random parameter space. Sampling in high-dimensional random space is very time-consuming. If the sampling of random space is conducted in full random space through stochastic collocation method, then the number of samples increases sharply with respect to the dimension of the random space. This is the notorious *curse of dimensionality*, which poses great difficulties for the stochastic approximation in the high-dimensional stochastic space. Dimension reduction techniques such as high dimensional model representation (HDMR) [31], proper orthogonal decomposition and principal component analysis, are very helpful to alleviate the difficulty. Among these methods, HDMR is one of the efficient stochastic dimension reduction techniques and has caught extensive attention of researchers in recent years [25, 26, 28, 34]. The HDMR was originally developed in the application of chemical models by [31] and now becomes a general set of quantitative assessment and analysis tools for capturing the high-dimensional relationship between model inputs and model outputs. HDMR has been used for improving the efficiency of deducing high-dimensional input-output system behavior, and can be employed to relieve the computation effort. Instead of dealing with the full high-dimensional random space, truncated HDMR technique can decompose a high-dimensional model into a set of low-dimensional models, each of which needs much less computational effort. The *curse of dimensionality* can be considerably suppressed by using HDMR. However, there exist some drawbacks in

traditional truncated HDMR [8, 31] and adaptive HDMR [25, 34]. For example, too many low-dimensional models need to be computed, and high-order of cooperative effects from important random variables are neglected. To overcome these drawbacks, we propose a hybrid HDMR, which implicitly uses complete HDMR in a moderate-dimensional space and explicitly uses the first-order truncated HDMR for the remaining dimensions. Thus the hybrid HDMR decomposes a high-dimensional model into a moderate-dimensional model and a few one-dimensional models. The moderate-dimensional space is spanned by the most active random dimensions. We use sensitivity analysis to identify the most active random dimensions. Then each low-dimensional stochastic model in the hybrid HDMR is solved by using sparse grid stochastic collocation method. The proposed hybrid HDMR renders a good approximation in stochastic space and substantially improves efficiency for high-dimensional stochastic models. Therefore, a good trade-off between computation complexity and dimension reduction error is achieved in the hybrid HDMR technique.

Porous media often have inherent complex heterogeneities and this poses a great challenge for numerical simulations. The heterogeneities may vary at very different scales. Simulating the flows in the porous media that use a very fine grid to capture the heterogeneities is computationally very expensive and possibly infeasible. However, disregarding the heterogeneities can lead to large errors. To take into the heterogeneities, many multiscale methods (see [2, 5, 6, 17] and references there) have been developed in recent twenty years and successfully used to simulate the flows in heterogeneous porous media. Among these multiscale methods, MMsFEM [6, 1] has become an important numerical multiscale method. The main idea of MMsFEM is to incorporate the small-scale information into finite element velocity basis functions and capture their effect on the large scale via mixed finite element formulation. The MMsFEM retains local conservation of velocity flux and has been found to be useful for solving flow and transport equations in heterogeneous porous media. In many cases, the multiscale basis functions can be computed overhead and used repeatedly in subsequent computations with different source terms, boundary conditions. This leads to a large computational saving in simulating the flow and transport process where the flow equation needs to be solved many times dynamically. When porous media exhibit non-separable scales, some global information is needed for representing non-local effects (e.g., channel, fracture and shale barriers) and used to construct multiscale basis functions. Using global information can significantly improve accuracy and render agreement with geological realism [1, 2, 18, 19]. If the global information is incorporated into MMsFEM, we refer to the MMsFEM as global MMsFEM (G-MMsFEM). In the paper, we extend the central concept of the global information to MMsFEM based on the hybrid HDMR.

We incorporate hybrid HDMR, sparse grid stochastic collocation method and MMsFEM together and develop a stochastic dimension reduction MMsFEM. For the MMsFEM, we hierarchically utilize approximate global information, which is in tune with the different parts of the hybrid HDMR. This implies the cheap computation for the approximate global information. The stochastic dimension reduction MMsFEM is able to reduce a high-dimensional stochastic multiscale model in both stochastic space and spatial physical space. Compared with traditional truncated HDMR techniques, much better efficiency and very good accuracy are achieved in the hybrid HDMR. We carefully analyze the proposed stochastic dimension reduction MMsFEM and investigate its applications for flows in random and heterogeneous porous media. Some important statistical properties (e.g., mean and variance) of the outputs of the

stochastic flow models are discussed.

The rest of the paper is organized as follows. In Section 2, we briefly introduce the background on the flow and transport system in random porous media. In Section 3, we present a general framework of HDMR and propose the hybrid HDMR technique. Some theoretical results and computation complexity are also addressed in the section. Section 4 is devoted to presenting MMsFEM based on the hybrid HDMR. In Section 5, a few numerical examples are presented to demonstrate the performance of MMsFEM based on the hybrid HDMR. Finally, some conclusions and closing remarks are made.

## 2. Background and notations.

**2.1. Two-phase flow system and its stochastic parametrization.** Let  $D$  be a convex bounded domain in  $\mathbb{R}^d$  ( $d = 2, 3$ ) and  $(\Omega, F, P)$  be a probability space, where  $\Omega$  is the set of outcomes,  $F$  is the  $\sigma$ -algebra generated by  $\Omega$ , and  $P$  is a probability measure.

Let  $k(x, \omega)$  be a random permeability field. We consider two-phase flow and transportation in the random permeability field. Here the two phases is referred as water and oil, designated by subscripts  $w$  and  $o$ , respectively. The equations of two-phase flow and transport can be given (in the absence of gravity and capillary effects) by flow equations

$$- \operatorname{div}(\lambda(S)k(x, \omega)\nabla p) = q, \quad (2.1)$$

$$\frac{\partial S}{\partial t} + \operatorname{div}(uf_w(S)) = 0, \quad (2.2)$$

where the total mobility  $\lambda(S)$  is given by  $\lambda(S) = \lambda_w(S) + \lambda_o(S)$  and  $q$  is a source term. Here  $\lambda_w(S) = k_{rw}(S)/\mu_w$  and  $\lambda_o(S) = k_{ro}(S)/\mu_o$ , where  $\mu_o$  and  $\mu_w$  are viscosities of oil and water phases, respectively, and  $k_{rw}(S)$  and  $k_{ro}(S)$  are relative permeabilities of oil and water phases, respectively. Here  $f_w(S)$  is the fractional flow of water and given by  $f_w = \lambda_w/(\lambda_w + \lambda_o)$ . The equation (2.1) is the flow equation and the equation (2.2) is the transport (or saturation) equation. According to Darcy's law, the total velocity  $u$  in (2.2) is given by

$$u = u_w + u_o = -\lambda(S)k\nabla p. \quad (2.3)$$

A random field can be usually parameterized by Fourier type expansion such as expansion of Karhunen-Loève (ref.[24]), polynomial chaos and wavelet [22]. This often gives rise to an infinite-dimensional random space. For computation, we truncate such an expansion to approximate the random field. Then  $k(x, \omega)$  can be formally described by

$$k(x, \omega) \approx k(x, \theta_1(\omega), \theta_2(\omega), \dots, \theta_N(\omega)). \quad (2.4)$$

For example, if a random field is characterized by a covariance structure, then the random field can be approximated by a finite sum of uncorrelated random variables through truncated Karhunen-Loève expansion (KLE). To get accurate approximation, a large number  $N$  of random parameters in (2.4) are required. This leads to a deterministic model in a high-dimensional random parameter space. For simplicity of presentation, we make the following assumption in the paper,

$$k(x, \omega) = k(x, \Theta), \quad \text{where } \Theta := \Theta(\omega) = (\theta(\omega_1), \theta_2(\omega), \dots, \theta_N(\omega)) \in \mathbb{R}^N.$$

Let  $I^N \subset \mathbb{R}^N$  be the image of  $\Theta$ , i.e.,  $I^N = \Theta(\Omega)$ , and  $\rho(\Theta) = \prod_{i=1}^N \rho_i(\theta_i)$  be the joint probability function of  $(\theta_1, \dots, \theta_N)$ . By equations (2.1), (2.2), (2.3) and

parametrization of permeability field, we formulate the following stochastic two-phase flow system: find random fields  $p(x, \Theta, t)$ ,  $v(x, \Theta, t)$ ,  $S(x, \Theta, t) : D \times I^N \times (0, T] \rightarrow \mathbb{R}$  such that they almost surely (a.s) satisfy the following equations subject to initial condition and boundary condition

$$\begin{cases} \operatorname{div}(u) = q \\ u = -\lambda(S)k(x, \Theta)\nabla p \\ \frac{\partial S}{\partial t} + \operatorname{div}(uf_w(S)) = 0. \end{cases} \quad (2.5)$$

For simplicity, we will suppress the spatial variable  $x$  and temporal variable  $t$  in the rest of paper when no ambiguity occurs.

**2.2. Sparse grid stochastic collocation method.** For stochastic two-phase flow systems (2.5), the statistic properties (e.g., mean and variance) of solutions are of the quantities of interest. These properties may be obtained by first sampling the parameter random space using, for example, a Monte Carlo method or sparse grid collocation method, then solving the deterministic problems for the samples and analyzing the corresponding results to obtain the desired statistic quantities. The convergence of Monte Carlo methods is slow and a very large number of samples may be required, which leads to high computational cost. Instead, we use Smolyak sparse grid collocation method [32], where the Smolyak interpolant  $\mathcal{A}(N + \ell, N)$  ( $\ell \geq 1$ ) is a linear combination of tensor product interpolations with the property: only products with a relatively small number of nodes are used and the linear combination is chosen in such a way that an interpolation property for  $N = 1$  is preserved for  $N > 1$  [4]. In the notation  $\mathcal{A}(N + \ell, N)$ , the  $\ell$  represents the interpolation level. Sparse grid collocation method is known to have the same asymptotic accuracy as tensor product collocation method, while requiring many fewer collocation points as the parameter dimension increases [29].

Sparse grids have been successfully applied to stochastic collocation in many recent works (e.g., [25, 26, 19]). Based on Smolyak formula (ref.[4]), a set of collocation points  $\{\Theta^{(j)}\}_{j=1}^{N_c}$  in  $I^N$  are specially chosen, where  $N_c$  is the number of collocation points. With these chosen collocation points and corresponding weights  $\{w^{(j)}\}_{j=1}^{N_c}$ , the statistical properties of the solutions can be obtained. At each of the collocation points, the deterministic system (2.5) is solved and the output, for example,  $S(x, \Theta^{(j)}, t)$  is obtained. Then the mean of  $S(x, \Theta, t)$  can be estimated by

$$\begin{aligned} E[S(x, \Theta, t)] &= \int_{I^N} S(x, \Theta, t) \rho(\Theta) d\Theta \\ &\approx \int_{I^N} \mathcal{A}(N + \ell, N) [S(x, \Theta, t)] \rho(\Theta) d\Theta = \sum_{j=1}^{N_c} S(x, \Theta^{(j)}, t) w^{(j)}. \end{aligned}$$

Here the weights  $\{w^{(j)}\}_{j=1}^{N_c}$  are determined by the basis functions of  $\mathcal{A}(N + \ell, N)$  and joint probability function  $\rho(\Theta)$  (ref.[13]). Similarly, the variance of  $S(x, \Theta^{(j)}, t)$  can be obtained by

$$\begin{aligned} \operatorname{Var}[S(x, \Theta, t)] &= \int_{I^N} (S(x, \Theta, t) - E[S(x, \Theta, t)])^2 \rho(\Theta) d\Theta \\ &\approx \sum_{j=1}^{N_c} S^2(x, \Theta^{(j)}, t) w^{(j)} - E^2[S(x, \Theta, t)]. \end{aligned} \quad (2.6)$$

Let  $H(N + \ell, N)$  denote the number of collocation points for Smolyak sparse grid interpolation  $\mathcal{A}(N + \ell, N)$ . Then it follows that (see [4])

$$H(N + \ell, N) \approx \frac{2^\ell}{\ell!} N^\ell \quad \text{for } N \gg 1. \quad (2.7)$$

This implies that the number of collocation points algebraically increases with respect to the dimension  $N$ . We utilize Smolyak sparse grid collocation for the numerical computation. The stochastic approximation of the Smolyak sparse grid collocation method depends on the number of collocation points and the dimension  $N$  of the random parameter space. The convergence analysis in [29] implies that the convergence of Smolyak sparse grid collocation is exponential with respect to the number of Smolyak points, but depends on the parameter dimension  $N$ . This exponential convergence rate behaves algebraically for  $N \gg 1$ .

**3. High dimensional model representation.** The truncated KLE leads to the two-phase flow deterministic model (2.5) with a high-dimensional parameter  $\Theta$ . The most challenge part of solving such a high-dimensional stochastic system is to discretize the high-dimensional random parameter. There exist a few methods for the discretization of the random space [3, 9, 26, 33, 35]. Among these methods, sparse grid collocation method has been widely used and generates completely decoupled systems, each of which is the same size as the deterministic system. The sparse grid collocation method is usually very efficient in moderate-dimensional spaces. However, when the dimension of the random parameter is large and the approximation of stochastic space is conducted in the high-dimensional stochastic space, a large number of collocation points are required by (2.7) and we need to solve deterministic model (2.5) at all of the collocation points. Consequently, the high dimensionality will deteriorate the efficiency for the collocation method. To overcome the difficulty, we use high dimensional model representation (HDMR) to reduce the stochastic dimension and enhance the efficiency of simulation. By truncating HDMR, the high-dimensional model can be decomposed into a set of low-dimensional models. Hence the computation efforts can be significantly relieved [25, 34].

In the section, we present a general HDMR framework and propose a hybrid HDMR.

**3.1. A general HDMR framework.** In the section, we adopt the concept on decompositions of multivariate functions in [21] to present a general HDMR framework in terms of operator theory. Let  $\mathcal{F}$  be a linear space of real-valued functions  $f(\Theta)$  defined on a cube  $I^N$  and  $\Theta := (\theta_1, \theta_2, \dots, \theta_N) \in I^N$ . In the paper,  $f(\Theta)$  represents a relationship between the input random vector  $\Theta$  and an output (e.g., saturation, water-cut).

To introduce HDMR, we define a set of projection operators as follows.

**DEFINITION 3.1.** [21] Let  $\{P_j\}_{j=1}^N$  be a set of commuting projection operators on  $\mathcal{F}$  satisfying the following property:

$$P_j(f) = f \quad \text{if } f \text{ does not depend on } \theta_j, \text{ and } P_j(f) \text{ does not depend on } \theta_j. \quad (3.1)$$

Let  $\mathbf{u} \subseteq \{1, \dots, N\}$  be a subset and  $I$  the identity operator, we define

$$P_{\mathbf{u}} := \prod_{j \in \mathbf{u}} P_j, \quad \text{and} \quad P_{\emptyset} := I.$$

The property (3.1) implies that  $P_{1,\dots,N}(f)$  is a constant. The identity operator  $I$  can be decomposed through the above commuting projections  $P_j$  ( $j = 1, \dots, N$ ):

$$I = \Pi_{j=1}^N [P_j + (I - P_j)] = \sum_{\mathbf{u} \subseteq \{1, \dots, N\}} \left( \Pi_{j \in \mathbf{u}} (I - P_j) \right) \left( \Pi_{j \in \{1, \dots, N\} \setminus \mathbf{u}} P_j \right). \quad (3.2)$$

Then the equation (3.2) leads to a decomposition of  $f$  given by

$$f = \sum_{\mathbf{u} \subseteq \{1, \dots, N\}} f_{\mathbf{u}}, \quad \text{where} \quad f_{\mathbf{u}} := \left( \Pi_{j \in \mathbf{u}} (I - P_j) \right) \left( \Pi_{j \in \{1, \dots, N\} \setminus \mathbf{u}} P_j \right) f. \quad (3.3)$$

where  $f_{\mathbf{u}}$  depends only on variables with indices in  $\mathbf{u}$  and  $f_{\emptyset}$  is a constant. The equation (3.3) gives an abstract HDMR formulation.

Associated with projection  $P_{\mathbf{u}}$ , we define  $\hat{P}_{\mathbf{u}} := P_{\{1, \dots, N\} \setminus \mathbf{u}}$ . Then  $\hat{P}_{\mathbf{u}} f$  only depends on the variables with indices in  $\mathbf{u}$ . As a result, the component  $f_{\mathbf{u}}$  of  $f$  defined in (3.3) can be recursively computed and explicitly computed, respectively by (ref. [21])

$$f_{\mathbf{u}} = \hat{P}_{\mathbf{u}} f - \sum_{\mathbf{v} \subsetneq \mathbf{u}} f_{\mathbf{v}} \quad \text{and} \quad f_{\mathbf{u}} = \sum_{\mathbf{v} \subseteq \mathbf{u}} (-1)^{|\mathbf{u}| - |\mathbf{v}|} \hat{P}_{\mathbf{v}} f. \quad (3.4)$$

Next we specify the projection operator in (3.1). Let  $\mu$  be a measure on Borel subsets of  $I^N$  and a product measure with unit mass, i.e.,

$$d\mu(\Theta) := \Pi_{i=1}^N d\mu_i(\theta_i), \quad \int_I d\mu_i(\theta_i) = 1, \quad i = 1, \dots, N. \quad (3.5)$$

We assume that the functions in  $\mathcal{F}$  are integrable with respect to  $\mu$ . For any  $f \in \mathcal{F}$  and  $j \in \{1, \dots, N\}$ , we define

$$P_j f(\Theta) = \int_I f(\Theta) d\mu_j(\theta_j). \quad (3.6)$$

We can check that the projections in (3.6) satisfy the conditions in (3.1). Consequently, for any  $\mathbf{u} \subseteq \{1, \dots, N\}$

$$\hat{P}_{\mathbf{u}} f(\Theta) = \int_{I^{N-|\mathbf{u}|}} f(\Theta) \Pi_{j \notin \mathbf{u}} d\mu_j(\theta_j).$$

In particular,

$$\hat{P}_{\emptyset} f = \int_{I^N} f(\Theta) d\mu(\Theta).$$

The inner product  $\langle \cdot, \cdot \rangle$  on  $\mathcal{F}$  induced by the measure  $\mu$  is defined as follows:

$$\langle f, h \rangle := \int_{I^N} f(\Theta) h(\Theta) d\mu(\Theta), \quad f, h \in \mathcal{F}.$$

The norm  $\| \cdot \|_{\mathcal{F}}$  on  $\mathcal{F}$  is defined by  $\|f\|_{\mathcal{F}} := \langle f, f \rangle^{1/2}$ .

**DEFINITION 3.2.** Let  $\mathcal{F}_{\emptyset}$  and  $\mathcal{F}_{\mathbf{u}}$  ( $\emptyset \neq \mathbf{u} \subseteq \{1, \dots, N\}$ ) be defined as follows:

$$\mathcal{F}_{\emptyset} := \{f \in \mathcal{F} : f \text{ is constant } C\},$$

$$\mathcal{F}_{\mathbf{u}} := \{f \in \mathcal{F} : f = f_{\mathbf{u}} \text{ depends on the variables with indices in } \mathbf{u}, \text{ and } f \in \cap_{j \in \mathbf{u}} \text{Ker}(P_j)\}.$$



By the definition of the projection operator, we can obtain the following theorem, which gives a decomposition of  $\mathcal{F}$ .

**THEOREM 3.3.** [30] *Let  $\mathcal{F}_{\mathbf{u}}$  be defined in Definition 3.2 for any  $\mathbf{u} \subseteq \{1, \dots, N\}$ . Then*

- (1)  $\mathcal{F} = \oplus_{\mathbf{u} \subseteq \{1, \dots, N\}} \mathcal{F}_{\mathbf{u}}$ .
- (2)  $\mathcal{F}_{\mathbf{u}} := \left( \prod_{j \in \mathbf{u}} (I - P_j) \right) \hat{P}_{\mathbf{u}} \mathcal{F}$ .

For any  $\mathbf{u} \subseteq \{1, \dots, N\}$ , we define

$$M_{\mathbf{u}} = \left( \prod_{j \in \mathbf{u}} (I - P_j) \right) \hat{P}_{\mathbf{u}}.$$

Then we can show the operators  $M_{\mathbf{u}}$  are commutative projection operators and mutually orthogonal, i.e.,

$$M_{\mathbf{u}}^2 = M_{\mathbf{u}}, \quad M_{\mathbf{u}} M_{\mathbf{v}} = 0 \quad \text{for } \mathbf{u} \neq \mathbf{v}.$$

Theorem 3.3 gives the HDMR expansion of  $f$  by

$$f = \sum_{\mathbf{u} \subseteq \{1, \dots, N\}} f_{\mathbf{u}} = \sum_{\mathbf{u} \subseteq \{1, \dots, N\}} M_{\mathbf{u}} f, \quad \text{where } M_{\mathbf{u}} f \in \mathcal{F}_{\mathbf{u}}. \quad (3.7)$$

Theorem 3.3 also implies that

$$\|f\|_{\mathcal{F}}^2 = \sum_{\mathbf{u} \subseteq \{1, \dots, N\}} \|M_{\mathbf{u}} f\|_{\mathcal{F}}^2, \quad \hat{P}_{\emptyset}(M_{\mathbf{u}} f) = 0 \quad \text{for } \mathbf{u} \neq \emptyset.$$

Any set of commutative projectors  $\{M_{\mathbf{v}}\}$  generate a distributive lattice whose elements are obtained by all possible combinations (addition and multiplication) of the projectors in the set. The lattice has a unique maximal projection operator  $\mathcal{M}$ , which gives the algebraically best approximation to the functions in  $\mathcal{F}$  [15]. The range of the maximal project operator  $\mathcal{M}$  for the lattice is the union of the ranges of  $\{M_{\mathbf{v}}\}$ . Because the commutative projectors  $\{M_{\mathbf{v}}\}$  are mutually orthogonal here, the maximal project operator  $\mathcal{M}$  and the range  $\mathcal{F}_{\mathcal{M}}$  of  $\mathcal{M}$  have their explicit expressions as follows:

$$\mathcal{M} = \sum_{\mathbf{v}} M_{\mathbf{v}}, \quad \mathcal{F}_{\mathcal{M}} = \oplus_{\mathbf{v}} \mathcal{F}_{\mathbf{v}}.$$

As more orthogonal projectors are retained in the set, the resultant approximation by its maximal projection operator  $\mathcal{M}$  will become better.

By using the notion of maximal projection operator, we define the order of HDMR approximation. Consider the set of projectors  $\{M_{\mathbf{u}} : \mathbf{u} \subseteq \{1, \dots, N\}, |\mathbf{u}| \leq q\}$ , the best approximation of  $f$  is given by

$$f \approx \mathcal{M}_q f := \sum_{\substack{\mathbf{u} \subseteq \{1, \dots, N\} \\ |\mathbf{u}| \leq q}} M_{\mathbf{u}} f.$$

This gives the  $q$ -th order HDMR approximation of  $f$ . General speaking, higher order HDMR approximations are never worse than lower order HDMR approximation. Adding new orthogonal projectors into a sum of orthogonal projectors always produce



a new maximal projection operator with better accuracy. It is argued in [31] that a low order HDMR (e.g., up to the third-order) usually gives a good approximation for most practical problems.

If the measure  $\mu$  in (3.5) is taken to be the probability measure

$$d\mu(\Theta) = \rho(\Theta)d\Theta = \prod_{i=1}^N \rho_i(\theta_i)d\theta_i,$$

then the resulting HDMR defined in (3.7) is called ANOVA-HDMR [31]. Fixed a point  $\bar{\Theta} := (\bar{\theta}_1, \bar{\theta}_2, \dots, \bar{\theta}_N)$ . If the measure  $\mu$  in (3.5) is taken as the Dirac measure located at the point  $\Theta$ , i.e.,

$$d\mu(\Theta) = \prod_{i=1}^N \delta(\theta_i - \bar{\theta}_i)d\theta_i,$$

then the resulting HDMR is Cut-HDMR. The point  $\bar{\Theta}$  is so called cut-point or anchor point. In ANOVA-HDMR, the  $\hat{P}_{\mathbf{v}}f$  involves  $N - |\mathbf{v}|$  dimensional integration and the computation of the components of  $f$  is expensive. In Cut-HDMR, the  $\hat{P}_{\mathbf{v}}f$  only involve the evaluation in a  $N - |\mathbf{v}|$  dimensional space and the computation is cheap and straightforward. Because of this reason, we use Cut-HDMR for the analysis and computation in the paper.

REMARK 3.1. *Using 3.4 and induction method, it can be shown [21] that*

$$\mathcal{M}_q f = \sum_{j=0}^q \alpha_{q,j} \sum_{\substack{\mathbf{u} \subseteq \{1, \dots, N\} \\ |\mathbf{u}| \leq q}} \hat{P}_{\mathbf{u}} f, \quad \text{where } \alpha_{q,j} = (-1)^{q-j} \binom{N-1-j}{q-j}. \quad (3.8)$$

*This gives a straightforward calculation for  $\mathcal{M}_q f$ .*

**3.2. A hybrid HDMR.** In this subsection, we follow the idea of abstract HDMR and present HDMR in a straightforward way. For practical computation, a hybrid HDMR is proposed for approximation of HDMR.

By (3.3), the HDMR of  $f(\Theta)$  can be represented in the form

$$f(\Theta) = f_0 + \sum_{i=1}^n f_i(\theta_i) + \sum_{i < j} f_{ij}(\theta_i, \theta_j) + \dots + f_{12\dots n}(\theta_1, \theta_2, \dots, \theta_n). \quad (3.9)$$

Here  $f_0$  is the zeroth-order component denoting the mean effect of  $f(\Theta)$ . The first-order component  $f_i(\theta_i)$  represents the individual contribution of the input  $\theta_i$  and the second-order component  $f_{ij}(\theta_i, \theta_j)$  represents the cooperative effects of  $\theta_i$  and  $\theta_j$  and so on. For most realistic physical systems, low-order HDMR  $\mathcal{M}_q f$  (e.g.,  $q \leq 3$ ) may give a good approximation [31]. The  $q$ -th order truncated HDMR can be rewritten as

$$\mathcal{M}_q f = f_0 + \sum_{m=1}^q \sum_{i_1 < \dots < i_m} f_{i_1 \dots i_m}(\theta_{i_1} \dots \theta_{i_m}). \quad (3.10)$$

The  $\mathcal{M}_q f$  can approximate  $f(\Theta)$  through truncating the HDMR. The  $\mathcal{M}_q f$  consists of a large number of component terms for high-dimensional models, the computation for all the components may be costly. For many cases, some high-order cooperative effects can not be neglected in computation models, especially when a model strongly relies on a few dependent variables. Consequently, the traditional truncated HDMR in (3.10) may not give very good approximation. To eliminate or alleviate the above

drawbacks of traditional truncated HDMR, we propose a new truncated HDMR, which is termed by hybrid HDMR.

We use Fourier amplitude sensitivity test [8] to pick up the most active dimensions from the component of the high-dimensional random vector  $\Theta$ . Let  $\{f_j\}_{j=1}^N$  be the first-order components defined in Eq.(3.9) and  $\sigma^2$  denote variance. We may assume that  $\sigma^2(f_1) \geq \sigma^2(f_2) \geq \dots \geq \sigma^2(f_N)$ . Otherwise, we re-order the index set  $\{j\}_{j=1}^N$  by their monotonically decreasing values of variance. Alternatively, we can use their monotonically decreasing values of expectation,  $E[f_1] \geq E[f_2] \geq \dots \geq E[f_N]$ . In the paper, we restrict ourself to the variance sensitivity. We set a threshold constant with  $0 < \zeta < 1$  and find an optimal  $J$  such that

$$\sum_{j=1}^J \sigma^2(\psi_j) / \sum_{j=1}^N \sigma^2(\psi_j) \geq \zeta. \quad (3.11)$$

Then we define  $\{\theta_j\}_{j=1}^J$  to be the most  $J$  active dimensions. We note that  $\sigma^2(f_j)$  gives us information on the impact of  $\theta_j$  when it is acting alone on the output. It is clear that if the change of the random input  $\theta_j$  within its range lead to a significant change on the output, then  $\sigma^2(f_j)$  is large. Therefore, the criterion (3.11) is reasonable to identify the most active dimensions. According to practical situation, we can adjust the value of  $\zeta$  in Eq. (3.11) such that  $J$  is much less than  $N$ . We define an index set  $\mathcal{J} = \{1, 2, \dots, J\}$ . Then we define the hybrid HDMR as following,

$$\mathcal{M}_{\mathcal{J}}f := \sum_{\mathbf{v} \subseteq \mathcal{J}} M_{\mathbf{v}}f + \sum_{i \in \{1, \dots, N\} \setminus \mathcal{J}} M_i f. \quad (3.12)$$

Here the set of projectors  $\{M_{\mathbf{v}}, \mathbf{v} \subseteq \mathcal{J}\} \cup \{M_i, i \in \{1, \dots, N\} \setminus \mathcal{J}\}$  generate a lattice and its maximal project operator is  $\mathcal{M}_{\mathcal{J}}$ . By (3.12), the hybrid HDMR  $\mathcal{M}_{\mathcal{J}}f$  consists of two parts: the first part  $\sum_{\mathbf{v} \subseteq \mathcal{J}} M_{\mathbf{v}}f$  is the complete HDMR on the most active dimensions indexed in  $\mathcal{J}$ , the second part  $\sum_{i \in \{1, \dots, N\} \setminus \mathcal{J}} M_i f$  is the first-order truncated HDMR on the remaining dimensions  $\{1, \dots, N\} \setminus \mathcal{J}$ . By (3.4), we can rewrite equation (3.12) by

$$\mathcal{M}_{\mathcal{J}}f := \hat{P}_{\mathcal{J}}f + \sum_{i \in \{1, \dots, N\} \setminus \mathcal{J}} f_i(\theta_i). \quad (3.13)$$

We note that the operator  $\hat{P}_{\mathcal{J}}$  projects the  $N$ -variable function  $f$  to a function defined on the  $J$  the most active dimensions. The term  $\hat{P}_{\mathcal{J}}f$  gives the dominant contribution to  $\mathcal{M}_{\mathcal{J}}f$ . Any first-order components are usually important, these components are retained in  $\mathcal{M}_{\mathcal{J}}f$ . In Cut-HDMR, there is no error for the approximation  $\mathcal{M}_{\mathcal{J}}f$  of  $f(\Theta)$  whenever the point  $\Theta$  is located the  $J$ -th dimensional subvolume  $\{\Theta_{\mathcal{J}}, \bar{\theta}_{J+1}, \dots, \bar{\theta}_N\}$  across the cut-point  $\bar{\Theta}$ .

If we use traditional truncated HDMR to approximate the term  $\hat{P}_{\mathcal{J}}f$  in (3.13), then we can obtain the adaptive HDMR developed in [25, 34],

$$\mathcal{M}_{\mathcal{J},q}^{ad}f := \sum_{\mathbf{v} \subseteq \mathcal{J}, |\mathbf{v}| \leq q} M_{\mathbf{v}}(\hat{P}_{\mathcal{J}}f) + \sum_{i \in \{1, \dots, N\} \setminus \mathcal{J}} f_i(\theta_i). \quad (3.14)$$

For simplicity of notation, we will suppress  $q$  in  $\mathcal{M}_{\mathcal{J},q}^{ad}$  in the paper when the truncation order  $q$  is not emphasized. The following proposition gives the relationship among  $f$ ,  $\mathcal{M}_{\mathcal{J}}f$  and  $\mathcal{M}_{\mathcal{J}}^{ad}f$ .

PROPOSITION 3.4. *Let  $\mathcal{M}_{\mathcal{J}}f$  and  $\mathcal{M}_{\mathcal{J}}^{ad}f$  be defined in (3.13) and (3.14), respectively. Then*

$$\|f - \mathcal{M}_{\mathcal{J}}^{ad}f\|_{\mathcal{F}}^2 = \|f - \mathcal{M}_{\mathcal{J}}f\|_{\mathcal{F}}^2 + \|\mathcal{M}_{\mathcal{J}}f - \mathcal{M}_{\mathcal{J}}^{ad}f\|_{\mathcal{F}}^2. \quad (3.15)$$

*Proof.* We note the equality

$$f - \mathcal{M}_{\mathcal{J}}^{ad}f = (f - \mathcal{M}_{\mathcal{J}}f) + (\mathcal{M}_{\mathcal{J}}f - \mathcal{M}_{\mathcal{J}}^{ad}f). \quad (3.16)$$

Because the set of HDMR components in  $f - \mathcal{M}_{\mathcal{J}}f$  do not intersect with the set of HDMR components in  $\mathcal{M}_{\mathcal{J}}f - \mathcal{M}_{\mathcal{J}}^{ad}f$ , Theorem 3.3 implies that

$$\langle f - \mathcal{M}_{\mathcal{J}}f, \mathcal{M}_{\mathcal{J}}f - \mathcal{M}_{\mathcal{J}}^{ad}f \rangle = 0. \quad (3.17)$$

Then the equation (3.15) follows immediately by combining equation (3.16) and (3.17).  $\square$

The equation (3.15) implies that hybrid HDMR has better approximation than adaptive HDMR.

The mean of  $f$  can be computed by summing the mean of all HDMR components of  $f$  for both ANOVA-HDMR and Cut-HDMR. The variance of  $f$  is the sum of the variance of HDMR components of  $f$  for ANOVA-HDMR. However, the direct summation of variance of components of  $f$  may not equal to variance of  $f$  for Cut-HDMR. Consequently, we may want to have a truncated ANOVA-HDMR to approximate the variance of  $f$  for Cut-HDMR. If we directly derive a truncated ANOVA-HDMR, the computation is very costly and more expensive than the direct computation of the variance of  $f$  itself. This is because a truncated ANOVA-HDMR requires computing many high-dimensional integrations. To overcome the difficulty, we can use an efficient two-step approach to derive a hybrid ANOVA-HDMR through the hybrid Cut-HDMR.

Let  $\Theta_{\mathcal{J}}$  be the  $J$ -dim variable with indices in  $\mathcal{J}$ . Using the hybrid HDMR formulation (3.13), the hybrid Cut-HDMR  $f^{cut}(\theta)$  of  $f$  has the following form

$$f^{cut}(\Theta) = \hat{f}^{cut}(\Theta_{\mathcal{J}}) + \sum_{i \in \{1, \dots, N\} \setminus \mathcal{J}} f_i^{cut}(\theta_i), \quad \text{where } \hat{f}^{cut}(\Theta_{\mathcal{J}}) := \hat{\mathbf{P}}_{\mathcal{J}}f. \quad (3.18)$$

We use  $\mathcal{M}_{\mathcal{J}}$  to act on  $f^{cut}(\Theta)$  to get a hybrid ANOVA-HDMR  $f^{anova}(\theta)$  of  $f$ , which has the following form

$$f^{anova}(\Theta) := \mathcal{M}_{\mathcal{J}}f^{cut} = \hat{f}^{anova}(\Theta_{\mathcal{J}}) + \sum_{i \in \{1, \dots, N\} \setminus \mathcal{J}} f_i^{anova}(\theta_i). \quad (3.19)$$

Then we have the following theorem.

THEOREM 3.5. *Let  $f^{cut}(\Theta)$  and  $f^{anova}(\Theta)$  be defined in (3.18) and (3.19), respectively. Then*

$$E[f^{cut}] = E[f^{anova}] = E[\hat{f}^{cut}] + \sum_{i \in \{1, \dots, N\} \setminus \mathcal{J}} E[f_i^{cut}], \quad (3.20)$$

$$\text{Var}[f^{cut}] = \text{Var}[f^{anova}] = \text{Var}[\hat{f}^{cut}] + \sum_{i \in \{1, \dots, N\} \setminus \mathcal{J}} \text{Var}[f_i^{cut}]. \quad (3.21)$$

*Proof.* For the proof, we need to calculate the HDMR components in  $f^{anova}(\Theta)$ . By definition, we derive the term  $\hat{f}^{anova}(\Theta_{\mathcal{J}})$  as follows,

$$\begin{aligned}\hat{f}^{anova}(\Theta_{\mathcal{J}}) &= \hat{\mathbf{P}}_{\mathcal{J}} f^{cut} = \int_{I^{N-J}} f^{cut}(\Theta) \Pi_{i \notin \mathcal{J}} d\mu_i(\theta_i) \\ &= \int_{I^{N-J}} \hat{f}^{cut}(\Theta_{\mathcal{J}}) \Pi_{i \notin \mathcal{J}} d\mu_i(\theta_i) + \sum_{i \in \{1, \dots, N\} \setminus \mathcal{J}} \int_{I^{N-J}} f_i^{cut}(\theta_i) \Pi_{i \notin \mathcal{J}} d\mu_i(\theta_i) \\ &= \hat{f}^{cut}(\Theta_{\mathcal{J}}) + \sum_{i \in \{1, \dots, N\} \setminus \mathcal{J}} E[f_i^{cut}].\end{aligned}\tag{3.22}$$

We define  $f_0^{anova} = E[f^{cut}]$ . Then for any  $i \in \{1, \dots, N\} \setminus \mathcal{J}$ , we have

$$\begin{aligned}f_i^{anova}(\theta_i) &= \int_{I^{N-1}} f^{cut}(\Theta) \Pi_{\substack{s=1 \\ s \neq i}}^N d\mu_s(\theta_s) - f_0^{anova} \\ &= \int_{I^{N-1}} \hat{f}^{cut}(\Theta_{\mathcal{J}}) \Pi_{\substack{s=1 \\ s \neq i}}^N d\mu_s(\theta_s) + \sum_{j \in \{1, \dots, N\} \setminus \mathcal{J}} \int_{I^{N-1}} f_j^{cut}(\theta_j) \Pi_{\substack{s=1 \\ s \neq i}}^N d\mu_s(\theta_s) - f_0^{anova} \\ &= f_i^{cut}(\theta_i) + E[\hat{f}^{cut}] + \sum_{\substack{j \in \{1, \dots, N\} \setminus \mathcal{J} \\ j \neq i}} E[f_j^{cut}] - f_0^{anova} \\ &:= f_i^{cut}(\theta_i) + C_i.\end{aligned}\tag{3.23}$$

It is obvious that  $E[f_i^{anova}] = 0$  by (3.23). Consequently, it follows that by (3.22),

$$E[f^{anova}] = E[\hat{f}^{anova}] = E[\hat{f}^{cut}] + \sum_{i \in \{1, \dots, N\} \setminus \mathcal{J}} E[f_i^{cut}] = E[f^{cut}].$$

This proves the equality (3.20). For ANOVA-HDMR, we note that

$$\begin{aligned}\text{Var}[f^{anova}] &= \|f^{anova}\|_{\mathcal{F}}^2 = \|\hat{f}^{anova}\|_{\mathcal{F}}^2 + \sum_{i \in \{1, \dots, N\} \setminus \mathcal{J}} \|f_i^{anova}\|_{\mathcal{F}}^2 \\ &= \text{Var}[\hat{f}^{anova}] + \sum_{i \in \{1, \dots, N\} \setminus \mathcal{J}} \text{Var}[f_i^{anova}] \\ &= \text{Var}[\hat{f}^{cut}] + \sum_{i \in \{1, \dots, N\} \setminus \mathcal{J}} \text{Var}[f_i^{cut}],\end{aligned}\tag{3.24}$$

where equations (3.22) and (3.23) have been used in the last step. Further, by equations (3.22) and (3.23) we have

$$\begin{aligned}f^{cut}(\Theta) &= \hat{f}^{cut}(\Theta_{\mathcal{J}}) + \sum_{i \in \{1, \dots, N\} \setminus \mathcal{J}} f_i^{cut}(\theta_i) \\ &= \hat{f}^{anova}(\Theta_{\mathcal{J}}) - \sum_{i \in \{1, \dots, N\} \setminus \mathcal{J}} E[f_i^{cut}] + \sum_{i \in \{1, \dots, N\} \setminus \mathcal{J}} (f_i^{anova}(\theta_i) - C_i) \\ &= f^{anova}(\Theta) - \left( \sum_{i \in \{1, \dots, N\} \setminus \mathcal{J}} (E[f_i^{cut}] + C_i) \right).\end{aligned}\tag{3.25}$$

This implies that  $\text{Var}[f^{cut}] = \text{Var}[f^{anova}]$ . Hence the proof is completed.  $\square$

Theorem 3.5 implies that we can compute the mean and variance of  $f^{cut}$  by directly summing the means and variances of the components of  $f^{cut}$ . This can significantly reduce the complexity of computation. In addition, the derived hybrid ANOVA-HDMR  $f^{anova}$  is easily obtain by using the two-step approach through the the hybrid Cut-HDMR  $f^{cut}$ . Our further calculation shows that the traditional truncated Cut-HDMR (3.10) and adaptive Cut-HDMR (3.14) do not have these merits. We use sparse grid quadrature [13] to compute the mean and variance in the paper.

Compared with the traditional truncated HDMR and adaptive HDMR, the hybrid HDMR defined in Eq.(3.13) has the following advantages: (a) The hybrid HDMR consists of significantly fewer terms than the traditional truncated HDMR and adaptive HDMR, the total computational effort of the hybrid HDMR can be substantially reduced. (b) Since the hybrid HDMR inherently includes the all cooperative contribution from the most active dimensions, approximation accuracy is not worse (maybe better) in hybrid HDMR than the traditional truncated HDMR and adaptive HDMR; (c) According to Theorem 3.5, the computation of variance of hybrid HDMR is much more efficient.

**3.3. Analysis of computation complexity.** In the previous subsection, we have addressed the accuracy of the hybrid HDMR and made some comparisons with traditional truncated HDMR and adaptive HDMR. In this subsection, we discuss the computation efficiency for the various HDMR techniques when Smolyak sparse grid collocation is used. We remind that the HDMR means Cut-HDMR.

Let  $C(f, \ell)$  be the number of sparse grid collocation points with level  $\ell$  in full random parameter dimension space and  $C(\mathcal{M}_q f, \ell)$  the total number of sparse grid collocation points with level  $\ell$  in the traditional truncated HDMR  $\mathcal{M}_q f$ . We define  $C(\mathcal{M}_{\mathcal{J}} f, \ell)$  and  $C(\mathcal{M}_{\mathcal{J}, q}^{ad} f, \ell)$  in the similar way. We first consider the case  $\ell = 2$  and  $q = 2$  and calculate the total number of sparse grid collocation points for the different approaches. By using (2.7) and (3.8), we have for  $N \gg 1$

$$\begin{aligned} C(f, 2) &= H(N + 2, N) \approx 2N^2, \\ C(\mathcal{M}_2 f, 2) &= \sum_{j=1}^2 \binom{N}{j} H(j + 2, j) = 5N + 13 \binom{N}{2} \approx 13N^2, \\ C(\mathcal{M}_{\mathcal{J}} f, 2) &= H(J + 2, J) + (N - J)H(1 + 2, 1) \approx 2J^2 + 5(N - J), \\ C(\mathcal{M}_{\mathcal{J}, 2}^{ad} f, 2) &= \sum_{j=1}^2 \binom{J}{j} H(j + 2, j) + (N - J)H(1 + 2, 1) \approx 13J^2 + 5N. \end{aligned}$$

Consequently, if  $J \approx N/2$ , it follows immediately that

$$C(\mathcal{M}_{\mathcal{J}} f, 2) < C(f, 2) < C(\mathcal{M}_{\mathcal{J}, 2}^{ad} f, 2) < C(\mathcal{M}_2 f, 2).$$

This means that the hybrid HDMR  $\mathcal{M}_{\mathcal{J}} f$  requires the smallest number of collocation points when  $\ell = 2$  and  $q = 2$ . This case is our particular interest in the paper.

Next we consider some other cases. By some subtle calculation, it can be shown that if  $J \approx 30$  and  $N > (\frac{J}{3})^{1.5}$ , then

$$C(\mathcal{M}_{\mathcal{J}} f, 3) < C(\mathcal{M}_{\mathcal{J}, 2}^{ad} f, 3) < C(\mathcal{M}_2 f, 3) < C(f, 3).$$

This implies that the computation effort in the hybrid HDMR  $\mathcal{M}_{\mathcal{J}} f$  is the least for  $\ell = 3$  when the number of the most active dimensions is moderate for a high-dimensional problem. However, if the number  $J$  of the most active dimensions is

large (e.g.,  $J \gg 30$ ), we can show that

$$C(\mathcal{M}_{\mathcal{J},q}^{ad}f, \ell) < C(\mathcal{M}_q f, \ell) < C(\mathcal{M}_{\mathcal{J}} f, \ell) < C(f, \ell), \quad \text{where } \ell \geq 3 \text{ and } q \leq 3. \quad (3.26)$$

The relationship (3.26) tells us that adaptive HDMR  $\mathcal{M}_{\mathcal{J},q}^{ad}f$  may be the most efficient if higher level collocation is required and  $J$  is large as well.

For a high-dimensional stochastic model, if the number  $J$  of the most active dimensions is large and high level (e.g., level 3 and above) sparse grid collocation is required to approximate the term  $\hat{\mathbf{P}}_{\mathcal{J}}f$  in (3.13), using the adaptive HDMR (3.14) may improve efficiency by the comparison (3.26). The adaptive HDMR has been extensively considered in many recent papers [16, 25, 26, 34]. By our experience, we find that the number of the most active dimensions  $J$  is often less than  $\frac{N}{2}$  as  $N$  is large. If  $\hat{\mathbf{P}}_{\mathcal{J}}f$  is smooth with respect to  $\Theta_{\mathcal{J}}$ , low level sparse grid collocation (e.g., level 2) can usually provide accurate approximation. Many problems can fall into this situation and they are our special interest in the paper.

**3.4. Integrating HDMR and sparse grid collocation method.** As stated in Subsection 2.2, the sparse grid stochastic collocation method can reduce the stochastic two-phase flow system (2.5) into a set of deterministic two-phase flow systems. However, it suffers from *curse of dimensionality* with increasing the number of the stochastic dimensions. Integrating HDMR and sparse grid collocation method is a practical way to overcome this difficulty and can significantly enhance the efficiency.

Without loss of generality, we use saturation solution  $S(\Theta)$  as an example to present the technique for the hybrid Cut-HDMR. By using (3.13), (3.4) and Smolyak sparse grid interpolation, we have

$$\begin{aligned} S(\Theta) &\approx \mathcal{M}_{\mathcal{J}}S(\Theta) = \hat{S}(\Theta_{\mathcal{J}}) + \sum_{i \in \{1, \dots, N\} \setminus \mathcal{J}} S_i(\theta_i) \\ &= \hat{S}(\Theta_{\mathcal{J}}) + \sum_{i \in \{1, \dots, N\} \setminus \mathcal{J}} \hat{S}(\theta_i) - (N - J)S_0 \\ &\approx \mathcal{A}(J + \ell, J)[\hat{S}(\Theta_{\mathcal{J}})] + \sum_{i \in \{1, \dots, N\} \setminus \mathcal{J}} \mathcal{A}(1 + \ell, 1)[\hat{S}(\theta_i)] - (N - J)S_0, \end{aligned} \quad (3.27)$$

where  $\hat{S}(\Theta_{\mathcal{J}}) = \hat{\mathbf{P}}_{\mathcal{J}}S(\Theta)$ ,  $\hat{S}(\theta_i) = \hat{\mathbf{P}}_i S(\Theta)$  and  $S_0 = \hat{\mathbf{P}}_{\emptyset} S(\Theta) = S(\bar{\Theta})$ . Here  $\bar{\Theta}$  is the cut-point for a Cut-HDMR. The choice of cut-point may effect on the approximation of truncated Cut-HDMR. The study in [12] argued that an optimal choice of the cut-point is the center point of a sparse grid quadrature. In the paper, we will use such a cut-point for computation. Due to Theorem 3.5, the mean and variance of  $S$  can be approximated by

$$\begin{aligned} E[S(\Theta)] &\approx E\left[\mathcal{A}(J + \ell, J)[\hat{S}(\Theta_{\mathcal{J}})]\right] + \sum_{i \in \{1, \dots, N\} \setminus \mathcal{J}} E\left[\mathcal{A}(1 + \ell, 1)[\hat{S}(\theta_i)]\right] \\ &\quad - (N - J)S_0, \\ \text{Var}[S(\Theta)] &\approx \text{Var}\left[\mathcal{A}(J + \ell, J)[\hat{S}(\Theta_{\mathcal{J}})]\right] + \sum_{i \in \{1, \dots, N\} \setminus \mathcal{J}} \text{Var}\left[\mathcal{A}(1 + \ell, 1)[\hat{S}(\theta_i)]\right]. \end{aligned} \quad (3.28)$$

From (3.27), we find that the approximation error of the mean and variance comes from the two sources: truncated HDMR and Smolyak sparse grid quadrature. Let  $\{\Theta_{\mathcal{J}}^{(j)}, w^{(j)}\}_{j=1}^{N_J}$  be the set of pairs of collocation points and weights in the most active

dimension space  $I^J$  and  $\{\theta_i^{(k)}, w^{(k)}\}_{k=1}^{N_i}$  the set of pairs of collocation points and weights in  $I$ . We note that  $N_J = H(J + \ell, J)$  and  $N_i = H(1 + \ell, 1)$ . Then by (3.28), the mean of  $S$  can be computed by

$$E[S(\Theta)] \approx \sum_{j=1}^{N_J} \hat{S}(\Theta_{\mathcal{J}}^{(j)}) w^{(j)} + \sum_{i \in \{1, \dots, N\} \setminus \mathcal{J}} \sum_{k=1}^{N_i} \hat{S}(\theta_i^{(k)}) w^{(k)} - (N - J)S_0. \quad (3.29)$$

Similarly, we can compute  $\text{Var}[S(\Theta)]$  in terms of evaluations of  $\hat{S}(\Theta_{\mathcal{J}})$  and  $\hat{S}(\theta_i)$  at the collocation points.

Because each of terms in adaptive HDMR  $\mathcal{M}_{\mathcal{J},q}^{ad} S$  are usually correlated each other, the variance of  $\mathcal{M}_{\mathcal{J},q}^{ad} S$  is not equal to the summation of the variance of all terms in (3.14). Let  $\{\Theta_i^{(n)}, w^{(n)}\}_{n=1}^{N_c}$  (where  $N_c = H(N + \ell, N)$ ) be the set of pairs of collocation points and weights in full dimensions  $I^N$ . To compute the variance using adaptive HDMR, we need project each collocation point  $\Theta_i^{(n)}$  onto the components  $\{\Theta_{\mathbf{v}}^{(n)} : \mathbf{v} \subseteq \mathcal{J}, |\mathbf{v}| \leq q\}$  and  $\{\theta_i^{(n)} : i \in \{1, \dots, N\} \setminus \mathcal{J}\}$  and interpolate all terms in (3.14). Then we use (2.6) to calculate the variance. This process involves at most  $H(N + \ell, N) [N - J + \sum_{j=1}^q \binom{J}{j}]$  Smolyak sparse grid interpolations. The computation of these interpolations is usually very expensive when  $N$  and  $J$  are large. The computation of variance using adaptive HDMR is usually much more expensive than using hybrid HDMR. The numerical experiments in Section 5 will confirm the point.

**4. Mixed multiscale finite element method.** In Section 3.4, we have discussed integrating hybrid HDMR and sparse grid collocation method to reduce the computation complexity from high-dimensional stochastic spaces. The permeability field is often heterogeneous in porous media. It is necessary to use a numerical method to capture the heterogeneity. MMsFEM is one of such numerical methods and has been widely used in simulating flows in heterogeneous porous media [2, 19]. To simulate the two-phase flow system (2.5), it is necessary to retain local conservation for velocity (or flux). To this end, we use MMsFEM to solve the flow equation and obtain locally conservative velocity. Using MMsFEM coarsens the multiscale model in spatial space and can significantly enhance the computation efficiency.

Corresponding to the hybrid HDMR expansion of  $S(\Theta)$  in (3.27), i.e.,

$$\mathcal{M}_{\mathcal{J}} S(\Theta) = \hat{S}(\Theta_{\mathcal{J}}) + \sum_{i \in \{1, \dots, N\} \setminus \mathcal{J}} \hat{S}(\theta_i) - (N - J)S_0,$$

the velocity  $u(\Theta)$  in (2.5) also admits the same hybrid HDMR expansion as following

$$\mathcal{M}_{\mathcal{J}} u(\Theta) = \hat{u}(\Theta_{\mathcal{J}}) + \sum_{i \in \{1, \dots, N\} \setminus \mathcal{J}} \hat{u}(\theta_i) - (N - J)u_0, \quad (4.1)$$

where  $\hat{u}(\Theta_{\mathcal{J}}) = \hat{\mathcal{P}}_{\mathcal{J}} u(\Theta)$ ,  $\hat{u}(\theta_i) = \hat{\mathcal{P}}_i u(\Theta)$  and  $u_0 = u(\bar{\Theta})$ . We use  $\hat{u}(\Theta_{\mathcal{J}})$  to obtain  $\hat{S}(\Theta_{\mathcal{J}})$ ,  $\hat{u}(\theta_i)$  to obtain  $\hat{S}(\theta_i)$  and  $u_0$  to obtain  $S_0$ .

Without loss of generality, we may assume that the boundary condition in the flow equation of (3.27) is no flow boundary condition. Let  $\hat{k}(\Theta_{\mathcal{J}}) = \hat{\mathcal{P}}_{\mathcal{J}} k(\Theta)$  and  $\hat{k}(\theta_i) = \hat{\mathcal{P}}_i k(\Theta)$ . Then we can uniformly formulate the mixed formulation of equations of  $\hat{u}(\Theta_{\mathcal{J}})$  and  $\hat{u}(\theta_i)$  as following,

$$\begin{cases} (\lambda(\hat{S})\hat{k})^{-1} \hat{u} + \nabla \hat{p} = 0 & \text{in } D \\ \text{div}(\hat{u}) = q & \text{in } D \\ \hat{u} \cdot n = 0 & \text{on } \partial D. \end{cases} \quad (4.2)$$



Here  $\hat{u}(\Theta_{\mathcal{T}})$  and  $\hat{u}(\theta_i)$  are corresponding to the coefficients  $\hat{k}(\Theta_{\mathcal{T}})$  and  $\hat{k}(\theta_i)$ , respectively. Let  $k_0 = k(\bar{\Theta})$ . Then  $u_0$  is the solution to equation (4.2) if  $\hat{k}$  and  $\hat{S}$  are replaced by  $k_0$  and  $S_0$ , respectively.

The weak mixed formulation of (4.2) reads: seek  $(\hat{u}, \hat{p}) \in H_0(\text{div}, D) \times L^2(D)/R$  such that they satisfy the equation

$$\begin{cases} \left( (\lambda(\hat{S})\hat{k})^{-1} \hat{u}, v \right) - (\text{div}(v), \hat{p}) = 0 & \forall v \in H_0(\text{div}, D) \\ (\text{div}(\hat{u}), r) = (q, r) & \forall r \in L^2(D). \end{cases}$$

Let  $V_h \subset H_0(\text{div}, D)$  and  $Q_h \subset L^2(D)/R$  be the finite element spaces for velocity and pressure, respectively. Then the numerical mixed formulation of (4.2) is to find  $(\hat{u}_h, \hat{p}_h) \in V_h \times Q_h$  such that they satisfy

$$\begin{cases} \left( (\lambda(\hat{S})\hat{k})^{-1} \hat{u}_h, v_h \right) - (\text{div}(v_h), \hat{p}_h) = 0 & \forall v_h \in V_h \\ (\text{div}(\hat{u}_h), r_h) = (q, r_h) & \forall r_h \in Q_h. \end{cases} \quad (4.3)$$

We use MMsFEM for (4.3). It means that mixed finite element approximation is performed on coarse grid, where the multiscale basis functions are defined. In MMsFEM, we use piecewise constant basis functions on coarse grid for pressure. For the velocity, we define multiscale velocity basis functions. The degree of freedom of multiscale velocity basis function is defined on interface of coarse grid. Let  $e_i^K$  be a generic edge or face of the coarse block  $K$ . The velocity multiscale basis equation associated with  $e_i^K$  is defined by

$$\begin{cases} -\text{div}(\hat{k} \nabla w_i^K) = \frac{1}{|K|} & \text{in } K \\ -\hat{k} \nabla w_i^K \cdot n = \begin{cases} \frac{b_i^K \cdot n}{\int_{e_i^K} b_i^K \cdot n ds} & \text{on } e_i^K \\ 0 & \text{else.} \end{cases} \end{cases} \quad (4.4)$$

For local mixed MsFEM [6],  $b_i^K = n$ , the normal vector. If the media demonstrate strong non-local features including channels, fracture and shale barriers, some limited global information is needed to define the boundary condition  $b_i^K$  to improve accuracy of approximation [2, 18]. We will specify the boundary condition  $b_i^K$  for different parts in (4.1). Then  $\psi_i^K = -\hat{k} \nabla w_i^K$  defines multiscale velocity basis function associated to  $e_i^K$ , and the multiscale finite dimensional space for velocity is defined by

$$V_h = \bigoplus_{K,i} \psi_i^K.$$

It is well-known that using approximate single-phase global velocity information can considerably improve accuracy for multiscale simulation of two-phase flows [1, 2, 18, 19]. For the two-phase flow system (2.5), the single-phase global velocity  $u_{sg}(\Theta)$  solves the following equation

$$\begin{cases} (k(\Theta))^{-1} u_{sg} + \nabla p_{sg} = 0 & \text{in } D \\ \text{div}(u_{sg}) = q & \text{in } D \\ u_{sg} \cdot n = 0 & \text{on } \partial D. \end{cases} \quad (4.5)$$

By using hybrid HDMR and sparse grid interpolation, the  $u_{sg}(\Theta)$  admits the following approximation

$$\begin{aligned}
\mathcal{M}_{\mathcal{J}} u_{sg}(\Theta) &= \hat{u}_{sg}(\Theta_{\mathcal{J}}) + \sum_{i \in \{1, \dots, N\} \setminus \mathcal{J}} \hat{u}_{sg}(\theta_i) - (N - J)u_{sg,0} \\
&\approx \left\{ u_{sg,0} + \sum_{j \in \mathcal{J}} u_{sg,j}(\theta_j) \right\} + \sum_{i \in \{1, \dots, N\} \setminus \mathcal{J}} \hat{u}_{sg}(\theta_i) - (N - J)u_{sg,0} \\
&= \left\{ (1 - J)u_{sg,0} + \sum_{j \in \mathcal{J}} \hat{u}_{sg}(\theta_j) \right\} + \sum_{i \in \{1, \dots, N\} \setminus \mathcal{J}} \hat{u}_{sg}(\theta_i) - (N - J)u_{sg,0} \\
&\approx \left\{ (1 - J)u_{sg,0} + \sum_{j \in \mathcal{J}} \mathcal{A}(1 + \ell, 1) [\hat{u}_{sg}(\theta_j)] \right\} \\
&\quad + \sum_{i \in \{1, \dots, N\} \setminus \mathcal{J}} \mathcal{A}(1 + \ell', 1) [\hat{u}_{sg}(\theta_i)] - (N - J)u_{sg,0}
\end{aligned}$$

where  $\hat{u}_{sg}(\theta_j) = \hat{P}_j u_{sg}(\Theta)$  and  $\hat{u}_{sg}(\theta_i) = \hat{P}_i u_{sg}(\Theta)$ . Here the interpolation levels  $\ell$  and  $\ell'$  can be different. Because  $\{\theta_j\}_{j \in \mathcal{J}}$  are the most active dimensions, it is usually desirable that  $\ell \geq \ell'$ . Now we are ready to specifically describe the multiscale finite element space for  $\hat{u}(\Theta_{\mathcal{J}})$ ,  $\hat{u}(\theta_i)$  ( $i \in \{1, \dots, N\} \setminus \mathcal{J}$ ) and  $u_0$ .

- To construct the multiscale basis functions for  $\hat{u}(\Theta_{\mathcal{J}})$ , we take  $\hat{k} = \hat{k}(\Theta_{\mathcal{J}})$  in (4.4) and

$$b_i^K = b_i^K(\Theta_{\mathcal{J}}) = \left( (1 - J)u_{sg,0} + \sum_{j \in \mathcal{J}} \mathcal{A}(1 + \ell, 1) [\hat{u}_{sg}(\theta_j)] \right) \Big|_{e_i^K}. \quad (4.6)$$

The multiscale finite element space for  $\hat{u}(\Theta_{\mathcal{J}})$  is defined by

$$V_h(\Theta_{\mathcal{J}}) = \bigoplus_{K,i} \psi_i^K(\Theta_{\mathcal{J}}), \quad \text{where } \psi_i^K(\Theta_{\mathcal{J}}) = -\hat{k}(\Theta_{\mathcal{J}}) \nabla w_i^K(\Theta_{\mathcal{J}}). \quad (4.7)$$

- To construct the multiscale basis functions for  $\hat{u}(\theta_i)$  ( $i \in \{1, \dots, N\} \setminus \mathcal{J}$ ), we take  $\hat{k} = \hat{k}(\theta_i)$  in (4.4) and

$$b_i^K = b_i^K(\theta_i) = \left( \mathcal{A}(1 + \ell', 1) [\hat{u}_{sg}(\theta_i)] \right) \Big|_{e_i^K}. \quad (4.8)$$

The multiscale finite element space for  $\hat{u}(\theta_i)$  is defined by

$$V_h(\theta_i) = \bigoplus_{K,i} \psi_i^K(\theta_i), \quad \text{where } \psi_i^K(\theta_i) = -\hat{k}(\theta_i) \nabla w_i^K(\theta_i). \quad (4.9)$$

- To construct the multiscale basis functions for  $u_0$ , we replace  $\hat{k}$  by  $k(\bar{\Theta})$  in (4.4) and

$$b_i^K = u_{sg,0} \Big|_{e_i^K}.$$

The multiscale finite element space for  $u_0$  is defined by

$$V_h(\bar{\Theta}) = \bigoplus_{K,i} \psi_i^K(\bar{\Theta}), \quad \text{where } \psi_i^K(\bar{\Theta}) = -k(\bar{\Theta}) \nabla w_i^K(\bar{\Theta}).$$

For different parts of the hybrid HDMR expansion  $\mathcal{M}_{\mathcal{J}}u(\Theta)$  defined in (4.1), we use different boundary conditions for multiscale basis equations. This increases the hierarchies of multiscale basis functions, which are in tune with the sensitivity of random parameter dimensions.

By (3.29), to compute the moments of outputs we only need to construct the multiscale finite element space  $V_h(\Theta_{\mathcal{J}})$  at the set of collocation point  $\{\Theta_{\mathcal{J}}^{(j)}\}_{j=1}^{N_J}$ , where  $N_J = H(J + \ell, J)$ . For arbitrary  $\Theta_{\mathcal{J}} \in I^J$ , the boundary condition  $b_i^K(\Theta_{\mathcal{J}})$  is completely determined by  $u_{sg}(\theta_j^{(m)})$ , where  $j \in \mathcal{J}$  and  $m = 1, \dots, H(1 + \ell, 1)$ . Here  $\{\theta_j^{(m)}\}$  is a set of collocation points in  $I$ . Similarly, we also need to construct the multiscale finite element space  $V_h(\theta_i)$  at the set of 1-dimensional collocation points  $\{\theta_i^{(j')}\}$ , where  $i \in \{1, \dots, N\} \setminus \mathcal{J}$  and  $j' = 1, \dots, H(1 + \ell, 1)$ . For arbitrary  $\theta_i \in I$ , the boundary condition  $b_i^K(\theta_i)$  is completely determined by  $u_{sg}(\theta_i^{(m')})$ , where  $m' = 1, \dots, H(1 + \ell', 1)$  and  $\{\theta_j^{(m')}\}$  is a set of collocation points in  $I$ . These single-phase global velocity information  $\{u_{sg}(\theta_j^{(m)}), u_{sg}(\theta_i^{(m')})\}$  is pre-computed and can be repeatedly used in the whole stochastic flow simulations.

The convergence analysis for the mixed MsFEM using approximate global information can be found in [18, 19]. From the above discussion, here the approximate global information is completely determined by stochastic single-phase velocity contributed from individual random dimensions  $\theta_i$  ( $i = 1, \dots, N$ ). According to the analysis in [18], the resonance error between the coarse mesh size  $h$  and the effect of individual contributions of  $\theta_i$  ( $i = 1, \dots, N$ ) is negligible. This error is a major source using local MMsFEM [19]. However, the resonance error between the coarse mesh size  $h$  and the effect of cooperative contributions of  $\Theta$  is retained, but is usually small. This is our motivation to use limited global information determined in 1-dimensional random space. If the random permeability field has low variance and does not exhibit global features, we can use local mixed MsFEM for the stochastic simulation.

Define

$$\tilde{u}_h(x, \Theta) := \mathcal{A}(J + \ell, J)[\hat{u}_h(x, \Theta_{\mathcal{J}})] + \sum_{i \in \{1, \dots, N\} \setminus \mathcal{J}} \mathcal{A}(1 + \ell, 1)[\hat{u}_h(x, \theta_i)] - (N - J)u_{0,h}(x).$$

For velocity, we actually compute  $\tilde{u}_h(x, \Theta)$ , namely, the numerical solutions  $\hat{u}_h(x, \Theta_{\mathcal{J}})$  and  $\hat{u}_h(x, \theta_i)$  are evaluated at their collocation points. We define  $\tilde{u}(x, \Theta)$  in the same way. Now we analyze the total error between  $u(x, \Theta)$  and  $\tilde{u}_h(x, \Theta)$ . By triangle inequality,

$$\begin{aligned} & E[\|u(x, \Theta) - \tilde{u}_h(x, \Theta)\|_{L^2(D)}] \\ & \leq E[\|u(x, \Theta) - \mathcal{M}_{\mathcal{J}}u(x, \Theta)\|_{L^2(D)}] + E[\|\mathcal{M}_{\mathcal{J}}u(x, \Theta) - \tilde{u}(x, \Theta)\|_{L^2(D)}] \\ & \quad + E[\|\tilde{u}(x, \Theta) - \tilde{u}_h(x, \Theta)\|_{L^2(D)}] \\ & := \mathcal{E}_{hdmr} + \mathcal{E}_{col} + \mathcal{E}_{ms}, \end{aligned} \quad (4.10)$$

where  $\mathcal{E}_{hdmr}$  represents the error from hybrid HDMR,  $\mathcal{E}_{col}$  the error by sparse grid collocation and  $\mathcal{E}_{ms}$  the error of MMsFEM. The error  $\mathcal{E}_{hdmr}$  depends on the choice of the most active dimensions  $\mathcal{J}$ , the error  $\mathcal{E}_{col}$  depends on the number of collocation points and random dimensions, and the error  $\mathcal{E}_{ms}$  relies on the coarse mesh size  $h$  and the approximation of the boundary conditions (4.6) and (4.8) to  $\hat{u}_{sg}(\Theta_{\mathcal{J}})$  and  $\hat{u}_{sg}(\theta_i)$ , respectively. According to the works in [2, 18], if the boundary conditions in (4.6) and (4.8) are replaced by  $\hat{u}_{sg}(\Theta_{\mathcal{J}})$  and  $\hat{u}_{sg}(\theta_i)$ , respectively, then the error  $\mathcal{E}_{ms}$

only depends on coarse mesh size  $h$ . The error  $\mathcal{E}_{col}$  has been extensively discussed in many recent literatures, e.g., [3, 4, 33], and we focus on the errors  $\mathcal{E}_{hdmr}$  and  $\mathcal{E}_{ms}$  in the paper.

**5. Numerical results.** In this section, we present a few numerical results for two-phase flow in random porous media. The hybrid HDMR and MMsFEM are incorporated together to enhance simulation efficiency. If the random porous media have some non-local spatial features and high variance, we use limited global information to construct multiscale velocity information to achieve better accuracy. We will combine MMsFEM and different truncated HDMR techniques (hybrid HDMR  $\mathcal{M}_{\mathcal{J}}$  and adaptive HDMR  $\mathcal{M}_{\mathcal{J}}^{ad}$ ) for the flow simulations. The accuracy and efficiency of computation will be compared for the different methods.

For numerical simulation, we assume that  $k(x, \Theta)$  to be a logarithmic random field, i.e.,  $k(x, \Theta) := \exp(a(x, \Theta))$ . This assumption assures that  $k(x, \Theta)$  is positive and the well-posedness of flow equation in (2.5). Here  $a(x, \omega)$  is a stochastic field and can be characterized by its covariance function  $\text{cov}[a]: \bar{D} \times \bar{D} \rightarrow \mathbb{R}$  by

$$\text{cov}[a](x_1, x_2) := E[(a(x_1) - E[a(x_1)])(a(x_2) - E[a(x_2)])].$$

For the numerical experiments, we use a two point exponential covariance function for the stochastic field  $a$ ,

$$\text{cov}[a](x_1, y_1; x_2, y_2) = \sigma^2 \exp\left(-\frac{|x_1 - x_2|^2}{2l_x^2} - \frac{|y_1 - y_2|^2}{2l_y^2}\right), \quad (5.1)$$

where  $(x_i, y_i)$  ( $i = 1, 2$ ) is the spatial coordinate in 2D,  $\sigma^2$  is the variance of stochastic field  $a$ , and  $l_x$  and  $l_y$  denote the correlation length in the  $x$ - and  $y$ -direction, respectively. We will specify these parameters for the covariance function in numerical examples. Using Karhunen-Loève expansion (KLE) the random field  $a(x, \Theta)$  admits the following decomposition

$$a(x, \Theta) := E[a] + \sum_{i=1}^N \sqrt{\lambda_i} b_i(x) \theta_i(\omega), \quad (5.2)$$

where the random vector  $\Theta := (\theta_1, \theta_2, \dots, \theta_N) \in \mathbb{R}^N$  and the random variables  $\{\theta_i(\omega)\}_{i=1}^N$  are mutually orthogonal and have zero mean and unit variance. The effects of permeability field  $a(x, \Theta)$  with uniform, beta and Gaussian distributions on mean and standard deviation of output were discussed in [23], where the numerical results showed the similar peak values of standard deviation for the three different distributions. Therefore, in the section we assume that each  $\theta_i$  is i.i.d. uniform on  $[-1, 1]$ . This will not lose the main feature of the output uncertainty and hurt the performance of the proposed methods.

When MMsFEM is used, the fine grid is coarsened to form a uniform coarse grid. We solve the pressure equation on the coarse grid using MMsFEM and then reconstruct the fine-scale velocity field as a superposition of the multiscale basis functions. The reconstructed velocity field is used to solve the saturation equation with a finite volume method on the fine grid. We solve the two-phase flow system (2.5) by the classical IMPES (Implicit Pressure Explicit Saturation). The temporal discretisation is an implicit scheme, which is unconditionally stable but leads to a nonlinear system (Newton-Raphson iteration solves the nonlinear system). In all numerical simulations, mixed multiscale basis functions are constructed once at the beginning of

computation. In the discussion, we refer to the grid where multiscale basis functions are constructed as the coarse grid. The limited global information is computed on the fine grid.

We compare the saturation fields and water-cut data as a function of pore volume injected (PVI). The water-cut is defined as the fraction of water in the produced fluid and is given by  $q_w/q_t$ , where  $q_t = q_o + q_w$ , with  $q_o$  and  $q_w$  being the flow rates of oil and water at the production edge of the model. In particular,  $q_w = \int_{\partial D^{out}} f(S)u \cdot n ds$ ,  $q_t = \int_{\partial \Omega^{out}} u \cdot n ds$ , where  $\partial D^{out}$  is the out-flow boundary. Pore volume injected is defined as  $PVI = \frac{1}{V_p} \int_0^t q_t(\tau) d\tau$ , with  $V_p$  being the total pore volume of the system, provides the dimensionless time for the displacement. We consider a traditional quarter five-spot problem (e.g., [1]) on a square domain  $D$ , where the water is injected at left top corner and oil is produced at the right lower corner of the rectangular domain (in the 2D examples).

**5.1. Linear transportation.** In this subsection, we consider the case when the transport (saturation) equation is linear in the flow system (2.5). For this purpose, we consider the mobility of water and mobility of oil defined as following,

$$\lambda_w(S) = S, \quad \lambda_o(S) = 1 - S.$$

Consequently, the total mobility  $\lambda(S) = 1$  and the fractional flow of water  $f_w(S) = S$ . Then the two-phase flow system (2.5) reduces to a linear flow model. Since  $\lambda(S) = 1$ , the velocity is not updated to compute saturation. We use the example to investigate the performance of the hybrid HDMR technique and MMsFEM for linear models.

The stochastic permeability field  $k(x, \Theta)$  is given by  $k(x, \Theta) = \exp(a(x, \Theta))$ , where  $a(x, \Theta)$  is characterized by the covariance function  $\text{cov}[a]$  defined in (5.1), whose parameters is defined as following:  $\sigma^2 = 1$ ,  $l_x = l_y = 0.1$ . We truncate the KLE (5.2) after the first 80 terms to represent the random field  $a(x, \Theta)$  with  $E[a] = 1$ . This implies that  $k(x, \Theta)$  is defined in an 80-dimensional random parameter space, i.e.,  $N = 80$ . Fig. 5.1 depicts a realization of the random field  $a(x, \Theta)$ . The stochastic field  $k(x, \Theta)$  is defined in a  $60 \times 60$  fine grid. We choose  $6 \times 6$  coarse grid and apply MMsFEM to compute velocity. The time step is taken to be 0.02 PVI for discretizing temporal variable.

To make comparison, we use various methods to simulate the linear two-phase flow system: fine-scale mixed finite element method (MFEM) on full random dimensions, MMsFEM on full random dimensions, MMsFEM based on adaptive HDMR, and MMsFEM based on hybrid HDMR. The solution computed by fine-scale MFEM is referred to as the reference solution in the paper. To assess the performance of MMsFEM and HDMR, we compute the mean and standard deviation (std) for the quantifies of interest from the two-phase flow model, such as saturation and water-cut. Since the random field in the example does not have strong non-local features, the local MMsFEM (L-MMsFEM) generally gives an accurate approximation [18]. Hence the local MMsFEM is utilized for simulation in this example. To choose the most active random dimensions for adaptive HDMR and hybrid HDMR, we choose a threshold constant  $\zeta = 0.9$  and use the criterion (3.11). This produces 31 most active dimensions. Smolyak sparse grid collocation method with level 2 is used to tackle stochastic space. We evaluate the flow model's outputs at the collocation points and use the associate weights to compute the mean and standard deviation of the outputs. Table 5.1 lists the number of deterministic models to be solved for the above four different methods. From the table, we find that MFEM on full random dimensions needs

to compute the largest number of deterministic models (fine-scale models), MMsFEM based on hybrid HDMR needs to compute the smallest number of deterministic models (coarse-scale models). This means that MMsFEM based on hybrid HDMR can significantly reduce the computation complexity.

TABLE 5.1  
number of deterministic models for the different methods

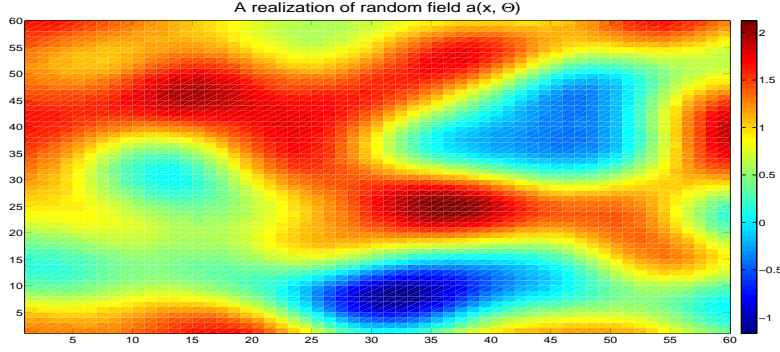
methods	number of deterministic models
MFEM on full random dim.	12961 fine-scale models
MMsFEM on full random dim	12961 coarse-scale models
MMsFEM based on adaptive HDMR	6446 coarse-scale models
MMsFEM based on hybrid HDMR	2231 coarse-scale models

Fig. 5.2 shows the point-wise mean of saturation map at PVI = 0.4. From the figure, we observe that both L-MMsFEM based on hybrid HDMR and L-MMsFEM based on adaptive HDMR provide very accurate approximation for the mean of saturation. They are almost identical to the reference mean computed by fine-scale MMFEM on full random dimensions. Fig. 5.3 illustrates the point-wise standard deviation of saturation at PVI = 0.4. By the figure, we have four observations: (1) Compared with reference standard deviation, the approximations of standard deviation by the three multiscale models (L-MMsFEM on full random dim., L-MMsFEM based on hybrid HDMR and L-MMsFEM based on adaptive HDMR) are good; (2) L-MMsFEM on full random dim. renders better approximation for standard deviation than L-MMsFEM based on truncated HDMR techniques; (3) L-MMsFEM based on hybrid HDMR gives almost the same standard deviation as the L-MMsFEM based on adaptive HDMR; (4) The variance mostly occurs around the front of water flow. Fig. 5.4 depicts the mean of water-cut curves. We see that the four water-cut curves are almost identical. This demonstrates that a good approximation has been achieved. Fig. 5.5 depicts the standard deviation of water-cut curves by the four different methods. By the figure, we see that both L-MMsFEM based on hybrid HDMR and L-MMsFEM based on adaptive HDMR can render a very good approximation for standard deviation at most time instances. The standard deviation of water-cut curve by L-MMsFEM based on hybrid HDMR is almost identical to the standard deviation of water-cut curve by L-MMsFEM based on adaptive HDMR.

Next we discuss the relative errors of mean and standard deviation in the sense of integration with regarding to physical domain. Let  $S_r/W_r$ ,  $S_{ms,f}/W_{ms,f}$ ,  $S_{ms,aH}/W_{ms,aH}$  and  $S_{ms,hH}/W_{ms,hH}$  be the saturation/water-cut using MFEM on full random dimensions, MMsFEM on full random dimensions, MMsFEM based on adaptive HDMR and MMsFEM based hybrid HDMR, respectively. Then we define the relative errors of saturation mean and saturation standard deviation between  $S_r$  and  $S_{ms,f}$  as following,

$$\mathcal{E}_{ms}^m(S) = \frac{\|E[S_r](x) - E[S_{ms,f}](x)\|_{L^1(D)}}{\|E[S_r](x)\|_{L^1(D)}}, \quad \mathcal{E}_{ms}^{std}(S) = \frac{\|\sigma[S_r](x) - \sigma[S_{ms,f}](x)\|_{L^1(D)}}{\|\sigma[S_r](x)\|_{L^1(D)}}, \quad (5.3)$$

where  $\sigma$  denotes the standard deviation operation. We can similarly define the relative errors  $\mathcal{E}_{aH}^m(S)$  and  $\mathcal{E}_{aH}^{std}(S)$  between  $S_r$  and  $S_{ms,aH}$ , and the relative errors  $\mathcal{E}_{hH}^m(S)$  and  $\mathcal{E}_{hH}^{std}(S)$  between  $S_r$  and  $S_{ms,hH}$ . We list these errors on saturation at 0.4 PVI in Table 5.2. From the table, we see that L-MMsFEM based on hybrid HDMR gives the same accuracy of mean as L-MMsFEM based on adaptive HDMR. They are both very

FIG. 5.1. A realization of random field  $a(x, \Theta)$ 

close to the error of mean by L-MMsFEM on full random dimensions. The relative error of standard deviation by L-MMsFEM based on hybrid HDMR also has a good agreement with that of L-MMsFEM based on adaptive HDMR. The table implies that the errors only from adaptive HDMR and hybrid HDMR are very small. The main source of errors is from spatial multiscale approximation.

TABLE 5.2  
relative errors of mean and std. on saturation (at 0.4 PVI)

methods	relative error of mean	relative error of std.
L-MMsFEM on full random dim.	1.223540e-002	5.302182e-002
L-MMsFEM based on adaptive HDMR	1.361910e-002	9.389291e-002
L-MMsFEM based on hybrid HDMR	1.361910e-002	9.482789e-002

For water-cut, the relative errors of mean and standard deviation between  $W_r$  and  $W_{ms,f}$  are defined by

$$\mathcal{E}_{ms}^m(W) = \frac{\|E[W_r](t) - E[W_{ms,f}](t)\|_{L^2([0,1])}}{\|E[W_r](t)\|_{L^2([0,1])}}, \quad \mathcal{E}_{ms}^{std}(W) = \frac{\|\sigma[W_r](t) - \sigma[W_{ms,f}](t)\|_{L^2([0,1])}}{\|\sigma[W_r](t)\|_{L^2([0,1])}}, \quad (5.4)$$

where  $[0, 1]$  is the temporal domain. The relative errors of water-cut such as  $\mathcal{E}_{ms}^m(W)$ ,  $\mathcal{E}_{aH}^{std}(W)$ ,  $\mathcal{E}_{hH}^m(W)$  and  $\mathcal{E}_{hH}^{std}(W)$ , are defined in a similar way. We list these relative errors of water-cut in Table 5.3. The table shows the similar behavior of these methods to the situation of saturation. The performance of L-MMsFEM based on hybrid HDMR is almost the same as the performance of L-MMsFEM based on adaptive HDMR.

TABLE 5.3  
relative errors of mean and std. on water-cut

methods	relative error of mean	relative error of std.
L-MMsFEM on full random dim.	1.445791e-002	4.339328e-002
L-MMsFEM based on adaptive HDMR	1.856908e-002	1.006136e-001
L-MMsFEM based on hybrid HDMR	1.856908e-002	1.040452e-001



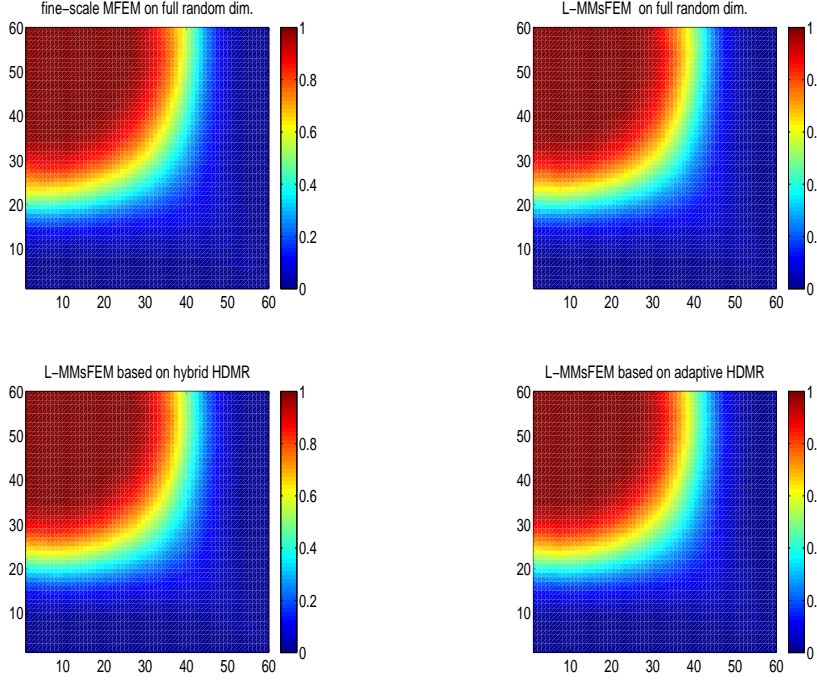


FIG. 5.2. Point-wise mean of saturation map at  $PVI=0.4$ . Top left: MFEM on full random dimensions; Top right: L-MMsFEM on full random dimensions; Bottom left: L-MMsFEM based on hybrid HDMR; Bottom right: L-MMsFEM based on adaptive HDMR

The number of the most active dimensions may have an important impact on computation efficiency and accuracy. To this end, we choose some different threshold constants in (3.11). Specifically, we take  $\zeta = 0.75, 0.8, 0.85, 0.9, 0.95$ . These threshold constants correspond to the number of the most active dimensions  $\dim(\Theta_{\mathcal{J}}) = 19, 22, 26, 31, 41$ . Fig. 5.6 illustrates the CUP times of L-MMsFEM based on adaptive HDMR and L-MMsFEM based on hybrid HDMR for the different number of the most active dimensions. From the figure, we get two observations: (1) CPU time of L-MMsFEM based on hybrid HDMR is only a fraction of the CUP time using L-MMsFEM based on adaptive HDMR; (2) As the number of the most active dimensions increases, the CPU time of L-MMsFEM based on hybrid HDMR increases mildly, but the CPU time of L-MMsFEM based on adaptive HDMR increases drastically. By the comparison, we see that hybrid HDMR is much more efficient than adaptive HDMR. There are two reasons for adaptive HDMR to take more CPU time: (1) adaptive HDMR needs to solve more deterministic models; (2) to compute standard deviation (or variance), a large number of stochastic interpolations are involved. Fig. 5.7 depicts the relative errors of mean (left) and standard deviation (right) for saturation at 0.4 PVI. By the figure, we find that increasing the number of the most active dimensions can substantially improve the accuracy of the mean and standard deviation for saturation. We also see that the L-MMsFEM based on hybrid HDMR has almost the same accuracy as the L-MMsFEM based on adaptive HDMR for sat-

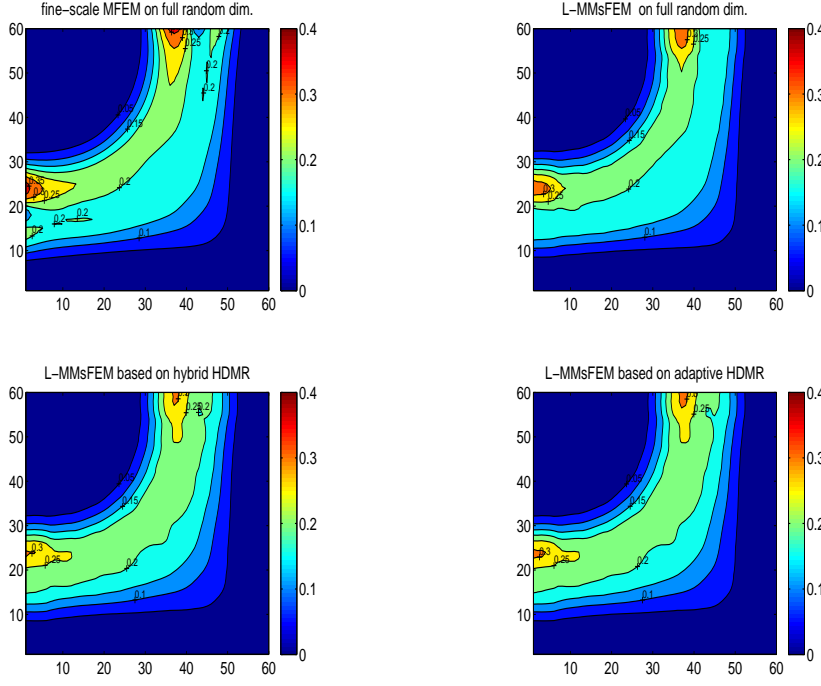


FIG. 5.3. Point-wise standard deviation of saturation contour at  $PVI=0.4$ . Top left: MFEM on full random dimensions; Top right: L-MMsFEM on full random dimensions; Bottom left: L-MMsFEM based on hybrid HDMR; Bottom right: L-MMsFEM based on adaptive HDMR

uration approximation. Fig. 5.8 describes the the relative errors of water-cut mean (left) and water-cut standard deviation (right) for the different number of the most active dimensions. From the figure, we observe that the effect of increasing  $\dim(\Theta_{\mathcal{J}})$  may not be as sensitive as the case of saturation for enhancement of accuracy. The reason is that the MMsFEM error is the main contribution of total error and has greater impact on the water-cut than HDMR itself.

**5.2. Non-linear transportation.** In the subsection, we consider the flow system (2.5) where the saturation is nonlinear. In the example, the mobility of water and the mobility of oil are defined by

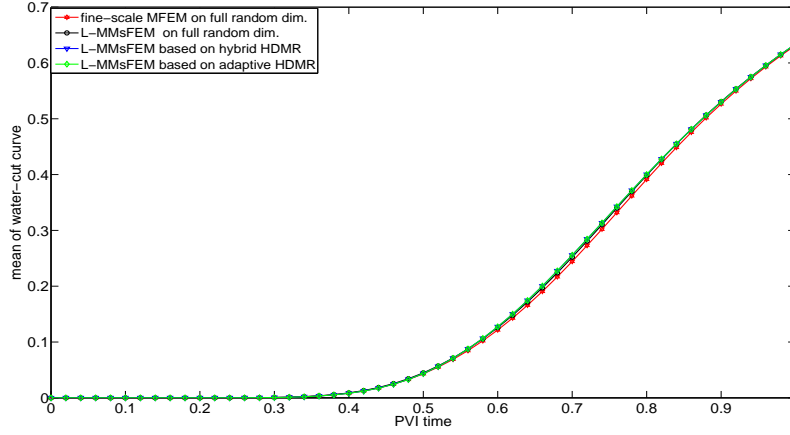
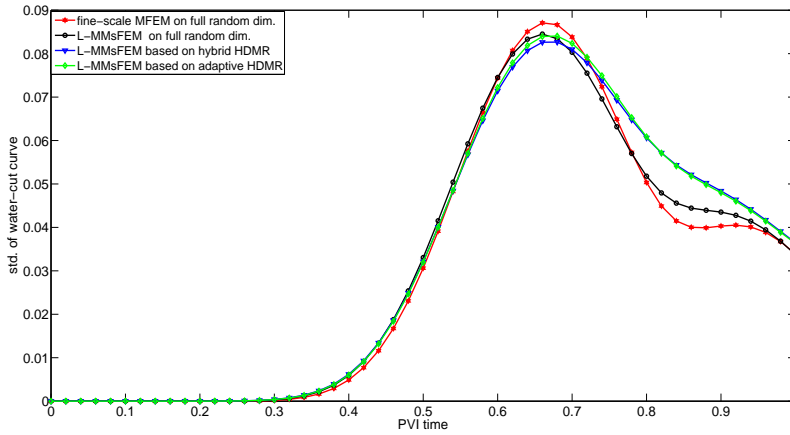
$$\lambda_w(S) = S^2/\mu_w, \quad \lambda_o(S) = (1 - S)^2/\mu_o.$$

Here  $\mu_w/\mu_o = 0.1$ , the ration between viscosity of water and oil. Consequently, the fractional function  $f_w(S)$  of water is given by

$$f_w(S) = \frac{S^2}{S^2 + 0.1(1 - S)^2}.$$

This results in a non-linear two-phase flow system.

We again consider the random permeability  $k(x, \Theta) = \exp(a(x, \Theta))$ . Here covariance function  $\text{cov}[a]$  of  $a(x, \Theta)$  is defined in (5.1) with  $l_x = l_y = 0.2$  and  $\sigma^2 = 1$ .

FIG. 5.4. *Mean of water-cut curves*FIG. 5.5. *Standard deviation of water-cut curves*

The mean of  $a(x, \Theta)$  is highly heterogeneous and its map is depicted in Fig. 5.9 (left). It is actually obtained by extracting and rescaling an SPE 10 [7] permeability field (the 45-th layer). We truncate the KLE (5.2) after the first 30 terms to represent the random field  $a(x, \Theta)$  and so the random field  $k(x, \Theta)$  is defined in an 30-dimensional random parameter space. Fig. 5.9 depicts a realization of the random field  $a(x, \Theta)$  (right). The stochastic field  $k(x, \Theta)$  is defined in a  $60 \times 60$  fine grid. We choose  $6 \times 6$  coarse grid for MMsFEM to compute velocity. To discretize the temporal variable of saturation equation, the time step is taken to be 0.01 PVI.

Fig. 5.9 shows that the permeability field  $k(x, \Theta)$  exhibit some channel features, which have important impacts on flows simulation. To achieve accurate simulation results, limited global information can improve the accuracy [1, 2, 18]. In the subsection we incorporate hybrid HDMR technique and adaptive HDMR technique with both local MMsFEM (L-MMsFEM) and global MMsFEM (G-MMsFEM) and develop dif-

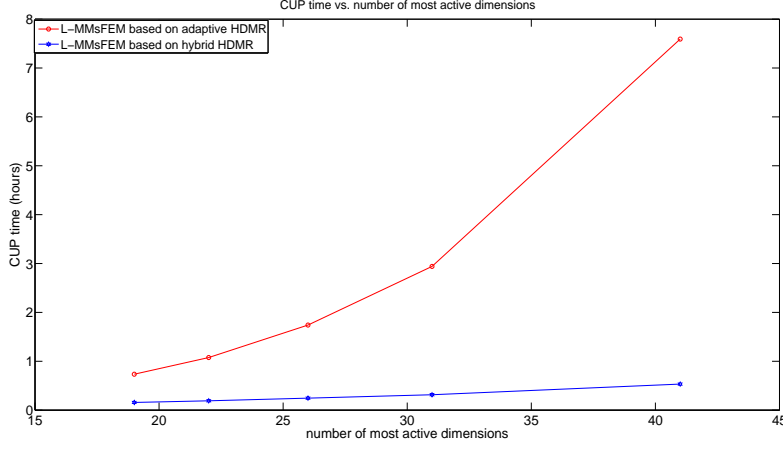


FIG. 5.6. CPU times for the different number of the most active dimensions, 19, 22, 26, 31, 41.

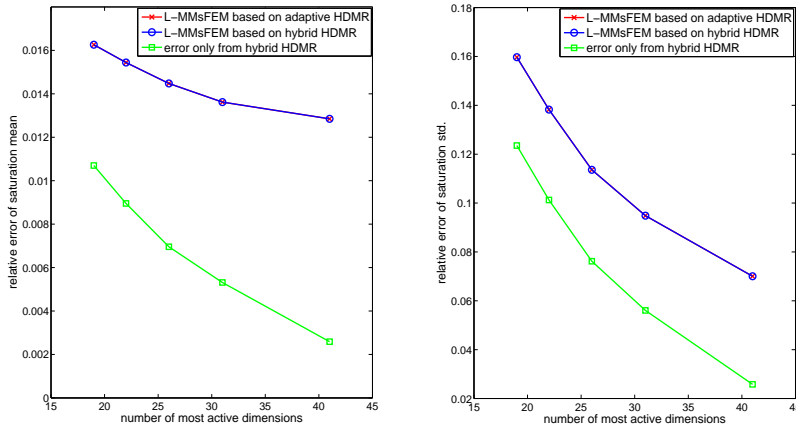


FIG. 5.7. Relative errors of mean (left) and standard deviation (right) for saturation at 0.4 PVI.

ferent methods. We will make some comparisons for these methods and explore their performance. To get the most active random dimensions for the truncated HDMR techniques, we choose a threshold constant  $\zeta = 0.9$  and use the criterion (3.11). This gives rise to 14 most active dimensions in the 30-dim random parameter space. We use the collocation points and weights associated with level 2 Smolyak sparse grid collocation method to compute mean and standard deviation.

Fig. 5.10 shows the point-wise mean of saturations at 0.4 PVI for different methods. We can clearly see that G-MMsFEM on full random dimensions has the best approximation to the reference solution given by fine-scale MFEM on full random dimensions. The figure also demonstrates that the multiscale methods with the truncated HDMR techniques provide good approximation. Fig. 5.11 describes the point-wise standard deviation of saturation at 0.4 PVI for those different methods. From the figure, we get two observations: (1) the large variance occurs around of front of

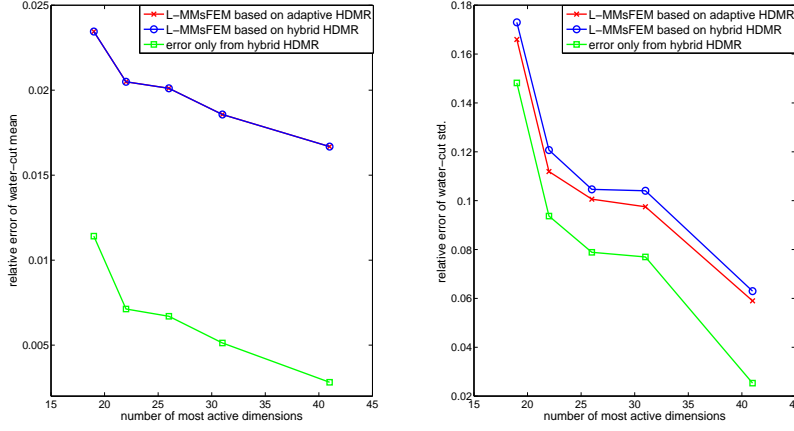
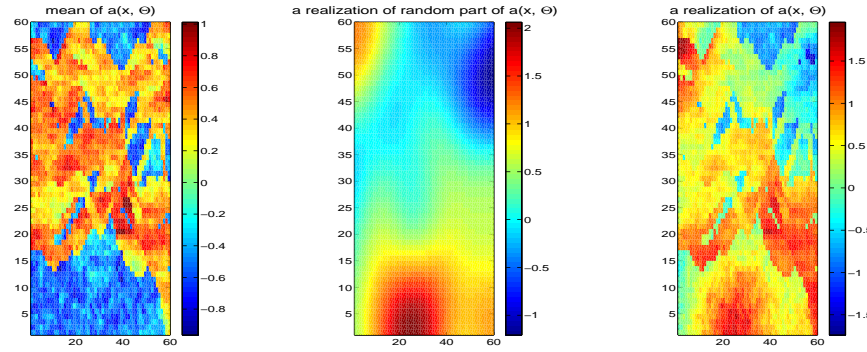
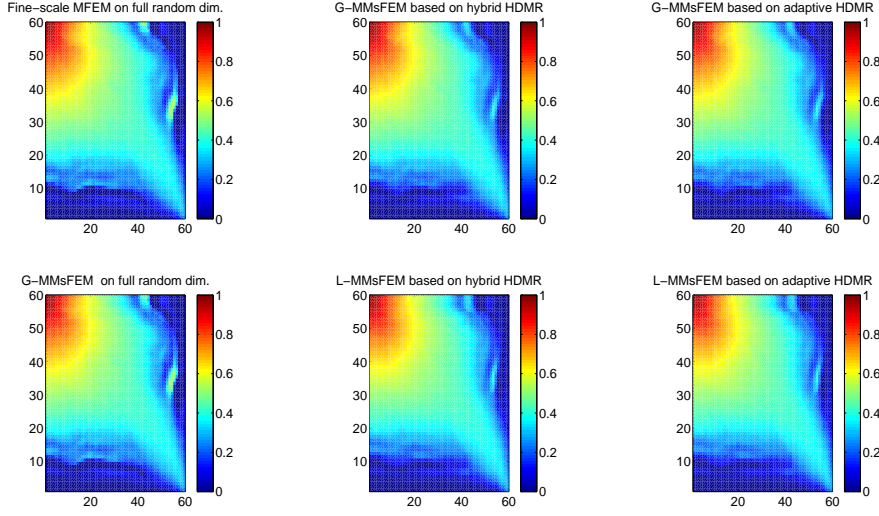
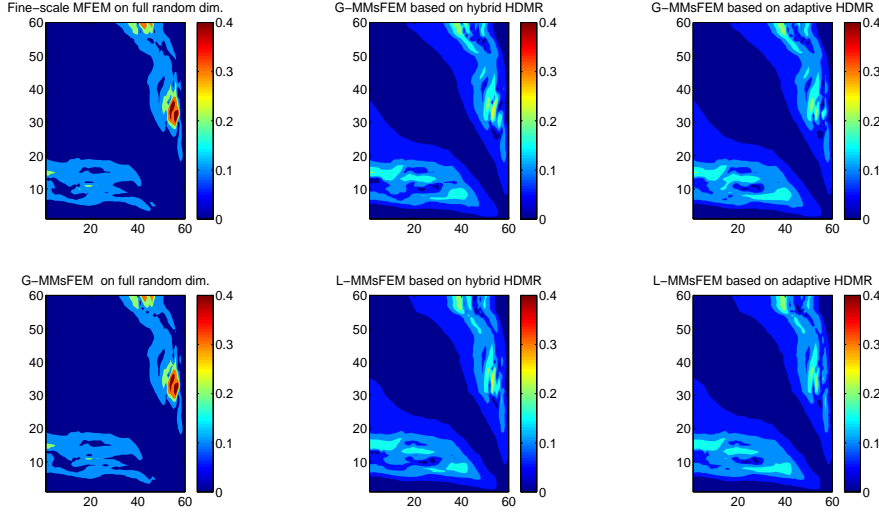


FIG. 5.8. Relative errors of mean (left) and standard deviation (right) for water-cut.

FIG. 5.9. A realization of random field  $a(x, \Theta)$ 

water flow; (2) the hybrid HDMR technique gives almost the same standard deviation map as the adaptive HDMR technique. We note that these behaviors share the similarities in Subsection 5.1. We also compute the water-cut curves for the different methods. For a better visualization, here we only present the water-cut curves by the four methods: MFEM on full random dim., G-MMsFEM on full random dim., G-MMsFEM based on adaptive HDMR and G-MMsFEM based on hybrid HDMR. Fig. 5.12 depicts the mean of water-cut curves for the four methods. From the figure, we see that the four curves overlap each other at almost all times. We note that there exists some fluctuation right after water break-through time. The reason may be that the value of water-cut changes sharply right after water break-through time. The standard deviation of the water-cut curves for the four methods are illustrated in Fig. 5.13. By the figure, we find that the variance of water-cut mostly occurs between the water-break-through time and time when water is full of the reservoir.

In order to carefully measure the differences caused by L-MMsFEM and G-MMsFEM integrating with adaptive HDMR and hybrid HDMR, we follow the concept in Subsection 5.1 and compute the relative errors between the reference solution

FIG. 5.10. *Point-wise mean of saturations at 0.4 PVI for different methods*FIG. 5.11. *Point-wise standard deviation of saturations at 0.4 PVI for different methods*

and the solutions of the various coarse models (MMsFEM with (or without) HDMR techniques). Here the reference solution is solved by fine-scale MFEM on full random dimensions. The relative errors of saturation and water-cut are defined in a similar way as in (5.3) and (5.3), respectively. Table 5.4 and 5.5 list the relative errors of mean and standard deviation for saturation at 0.4 PVI and water-cut, respectively. From the two tables, we find: (1) limited global information can enhance the accuracy of MMsFEM; (2) the results by adaptive and hybrid HDMR are very close to each

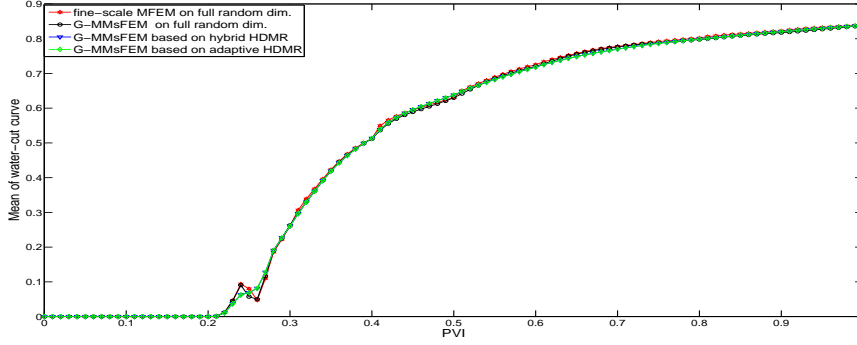


FIG. 5.12. Mean of water-cut curves for MFEM on full random dim., G-MMsFEM on full random dim., G-MMsFEM based on adaptive HDMR and G-MMsFEM based on hybrid HDMR

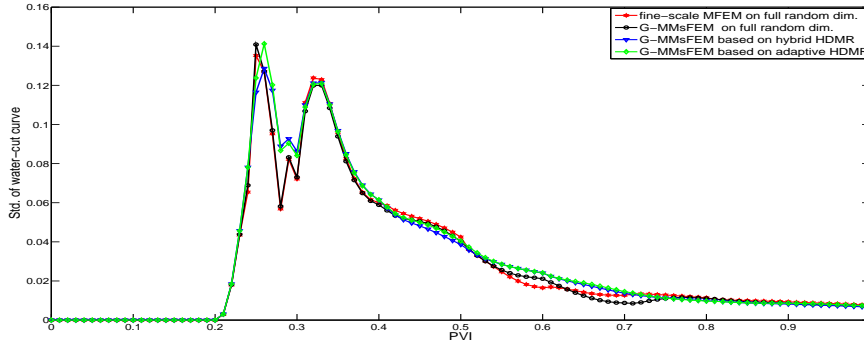


FIG. 5.13. Standard deviation of water-cut curves for MFEM on full random dim., G-MMsFEM on full random dim., G-MMsFEM based on adaptive HDMR and G-MMsFEM based on hybrid HDMR

other.

TABLE 5.4  
relative errors of mean and std. on saturation (at 0.4 PVI)

methods	relative error of mean	relative error of std.
L-MMsFEM on full random dim.	4.170726e-002	1.545560e-001
G-MMsFEM on full random dim.	8.950902e-003	4.027147e-002
L-MMsFEM based on adaptive HDMR	4.788991e-002	2.051337e-001
G-MMsFEM based on adaptive HDMR	3.283281e-002	1.608348e-001
L-MMsFEM based on hybrid HDMR	4.788991e-002	2.013145e-001
G-MMsFEM based on hybrid HDMR	3.283281e-002	1.520236e-001

Finally we discuss the efficiency for the different approaches. To the end, we record the CPU times for the seven different methods: fine-scale MFEM on full random dimensions, L-MMsFEM on full random dimensions, G-MMsFEM on full random dimensions, L-MMsFEM based on adaptive HDMR, G-MMsFEM based on



TABLE 5.5  
relative errors of mean and std. on water-cut

methods	relative error of mean	relative error of std.
L-MMsFEM on full random dim.	1.320466e-002	8.836354e-002
G-MMsFEM on full random dim.	8.000854e-003	4.298053e-002
L-MMsFEM based on adaptive HDMR	1.423242e-002	1.352706e-001
G-MMsFEM based on adaptive HDMR	1.162734e-002	1.212131e-001
L-MMsFEM based on hybrid HDMR	1.423242e-002	1.447519e-001
G-MMsFEM based on hybrid HDMR	1.162734e-002	1.248755e-001

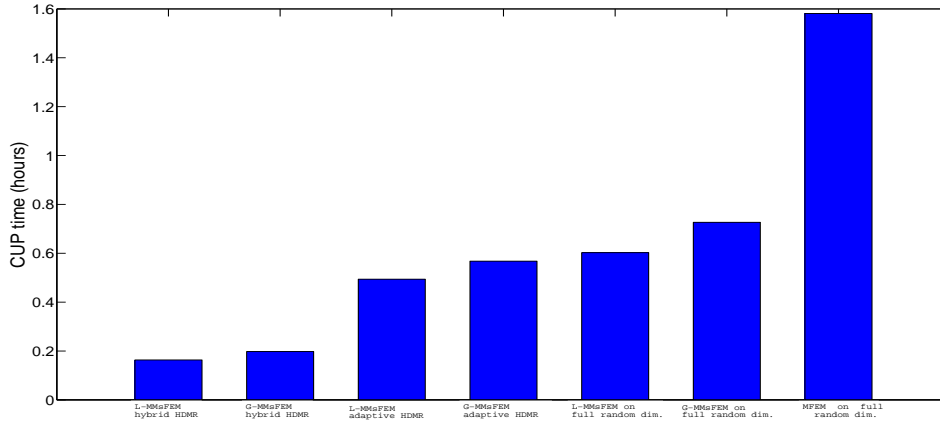


FIG. 5.14. CPU times for different approximation models

adaptive HDMR, L-MMsFEM based on hybrid HDMR and G-MMsFEM based on hybrid HDMR. Fig. 5.14 shows the CUP times for the seven different approaches. We see that the approaches of hybrid HDMR need the least CUP time and achieves the best efficiency. The CPU time of hybrid HDMR is only a fraction of the CUP time of adaptive HDMR. In the meanwhile, the CUP time of G-MMsFEM is slightly larger than the CUP time of L-MMsFEM when truncated HDMR techniques are synthesized. For adaptive HDMR, a lot of CUP time is spent in computing variance. However, computation of variance in hybrid HDMR is straightforward and requires very little CPU time. This shows the merit of the hybrid HDMR.

**6. Conclusions.** In the paper, we presented a framework of high-dimensional model representation (HDMR) for stochastic MMsFEM. A hybrid HDMR technique has been proposed and it decomposes a high-dimensional stochastic model into a moderate-dimensional stochastic model and a few one-dimensional stochastic models. Some criteria were presented to determine the moderate dimensions. We have developed MMsFEM based on hybrid HDMR by integrating it with MMsFEM. MMsFEM based on hybrid HDMR can significantly reduce the original model's complexity in both multiscale physical space and high-dimensional stochastic space. In the framework, we also presented MMsFEM based on adaptive HDMR, which has been widely used in stochastic model reduction. Compared with adaptive HDMR, the hybrid HDMR is much more efficient and retain the the same (or better) accuracy as adap-

tive HDMR. The computation time of the hybrid HDMR is a fraction of the computation time of adaptive HDMR. To capture strong non-local features in multiscale models, we have incorporated important global information into multiscale computation. This can improve the approximation accuracy of the proposed coarse multiscale models. We carefully analyzed the proposed MMsFEM using HDMR techniques and discussed computation efficiency and approximation errors. The MMsFEM based on HDMR techniques was applied to flows in random porous media. Both linear and non-linear stochastic flows models were discussed and the numerical results confirmed the performance of the proposed approaches.

## REFERENCES

- [1] J.E. Aarnes, *On the use of a mixed multiscale finite element method for greater flexibility and increased speed or improved accuracy in reservoir simulation*, Multiscale Model. Simul., 2 (2004), pp.421–439.
- [2] J. E. AARNES, Y. EFENDIEV AND L. JIANG, *Mixed multiscale finite element methods using limited global information*, Multiscale Model. Simul., 7 (2008), pp. 655–676.
- [3] I. BABUŠKA, F. NOBILE AND G. ZOURARIS, *A stochastic collocation method for elliptic partial differential equations with random input data*, SIAM J. Numer. Anal., 45 (2007), pp. 1005–1034.
- [4] V. BARTHELMANN, E. NOVAK AND K. RITTER, *High dimensional polynomial interpolation on sparse grids*, Advanced in Computational Mathematics 12 (2000), pp. 273–288.
- [5] Y. CHEN, L. J. DURLOFSKY, M. GERRITSEN, AND X. H. WEN, *A coupled local-global upscaling approach for simulating flow in highly heterogeneous formations*, Adv. in Water Res., 26 (2003), pp. 1041-1060.
- [6] Z. CHEN AND T. Y. HOU, *A mixed multiscale finite element method for elliptic problems with oscillating coefficients*, Math. Comp., 72 (2002), pp. 541–576.
- [7] M. CHRISTIE AND M. BLUNT, *Tenth SPE comparative solution project: A comparison of upscaling techniques*, SPE Reser. Eval. Eng., 4 (2001), pp. 308–317.
- [8] R.I. CUKIER, J.H. SCHAIBLY, AND K.E. SHULER, *Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. III. Analysis of the approximations*, Journal of Chemical Physics, 63 (1975), pp. 1140–1149.
- [9] G.S. FISHMAN, Monte Carlo, Concepts, Algorithms and Applications, Springer-Verlag, New York, 1996.
- [10] B. GANAPATHYSUBRAMANIAN AND N. ZABARAS, *A stochastic multiscale framework for modeling flow through random heterogeneous prous media*, J. Comput. Physics., 228 (2009), pp.591–618.
- [11] B. GANIS, I. YOTOV AND M. ZHONG, *A stochastic mortar mixed finite element method for flow in porous media with multiple rock types*, SIAM J. Sci. Comp., 33 (2011), pp. 1439–1474.
- [12] Z. GAO AND J. S. HESTHAVEN, *On ANOVA expansion and strategies for choosing the anchor point*, Applied Mathematics and Computation, 217 (2010), pp. 3274–3285.
- [13] T. GERSTNER AND M. GRIEBEL, *Numerical integration using sparse grids*, Numerical Algorithm 18 (1998), pp. 209–232.
- [14] V. GINTING, A. MALQVIST AND M. PRESCHO, *A novel method for solving multiscale elliptic problems with randomly perturbed data*, Multiscale Model. Simul., 8 (2010), pp. 977–996.
- [15] W. J. GORDON, *Distributive lattices and the approximation of multivariate functions*, in: *Proc. Symp. Approximation with Spectial Emphasis on Spline Functions* (edited by I.J. Schoenberg), Academic Press, New York, 1969.
- [16] M. GRIEBEL AND M. HOLTZ, *Dimension-wise integration of high-dimensional functions with applications to finance*, Journal of Complexity, 26 (2010), pp. 455–489.
- [17] P. JENNY, S. H. LEE, AND H. TCHELEPI, *Multi-scale finite volume method for elliptic problems in subsurface flow simulation*, J. Comput. Phys., 187 (2003), pp. 47–67.
- [18] L. JIANG, Y. EFENDIEV AND I. MISHEV, *Mixed multiscale finite element methods using approximate global information based on partial upscaling*, Comput. Geosci., 14 (2010), pp. 319–341.
- [19] L. JIANG, I. MISHEV, Y. LI, *Stochastic mixed multiscale finite element methods and their applications in reandom porous media*, Comput. Methods Appl. Mech. Engrg., 119(2011), pp. 2721–2740.
- [20] L. JIANG AND M. PRESCHO, *A resourceful splitting technique with applications to deterministic*

- and stochastic multiscale finite element methods, *Multiscale Model. Simul.*, 10 (2012), pp. 954–985.
- [21] F. Y. KUO, I. H. SLOAN, G. W. WASILKOWSKI, AND H. WOŹNIAKOWSKI, *On decompositions of multivariate functions*, *Math. Comp.*, 79 (2009), pp. 953–966.
  - [22] O. P. LEMAÎTRE AND O. M. KNIO, *Spectral methods for uncertainty quantification: with applications to computational fluid dynamics*, Springer–New York, 2010.
  - [23] G. LIN AND A. TARTAKOVSKY, *Numerical studies of three-dimensional stochastic Darcys equation and stochastic advection-diffusion-dispersion equation*, *Journal of Scientific Computing*, 43 (2010), pp. 92–117.
  - [24] M. LOËVE, *Probability Theory* (4th ed.), Springer Verlag, Berlin, 1977.
  - [25] X. MA AND N. ZABARAS, *A adaptive high-dimensional stochastic model representation technique for the solution of stochastic partial differential equations*, *J. Comput. Phys.*, 229 (2010), pp. 3884–3915.
  - [26] X. MA AND N. ZABARAS, *A stochastic mixed finite element heterogeneous multiscale method for flow in porous media*, *J. Comput. Phys.*, 230 (2011), pp. 4696–4722.
  - [27] H. G. MATTHIES AND A. KEESE *Galerkin methods for linear and nonlinear elliptic stochastic partial differential equations*, *Comput. Methods Appl. Mech. Engrg.*, 194 (2005), pp.1295–1331.
  - [28] N.C. NGUYEN, *A multiscale reduced-basis method for parameterized elliptic partial differential equations with multiple scales*, *J. Comput. Phys.*, 227 (2008), pp. 9807–9822.
  - [29] F. NOBILE, R. TEMPONE AND C. G. WEBSTER, *A sparse grid stochastic collocation method for partial differential equations with random input data*, *SIAM J. Numer. Anal.*, 46 (2008), pp. 2309–2345.
  - [30] H. RABITZ, Ö. F. ALLIS, *General foundations of high-dimensional model representations*, *Journal of Mathematical Chemistry* 25 (1999), pp. 197–233.
  - [31] H. RABITZ, Ö. F. ALLIS, J. SHORTER AND K. SHIM, *Efficient input-output model representation*, *Comput. Phys. Commun.*, 117 (1999), pp. 11–20.
  - [32] S. SMOLYAK, *Quadrature and interpolation formulas for tensor products of certain classes of functions*, *Soviet Math. Dokl.* 4, 1963, pp. 240–243.
  - [33] D. XIU AND JAN S. HESTHAVEN, *High-order collocation methods for differential equations with random inputs*, *SIAM J. Sci. Comput.* 27 (2005), pp. 1118–1139.
  - [34] X. YANG, M. CHOI, G. LIN AND G. E. KARNIADAKIS, *Adaptive ANOVA decomposition of stochastic incompressible and compressible flows*, *J. Comput. Phys.*, 2011, doi:10.1016/j.jcp.2011.10.028.
  - [35] D. ZHANG, L. LI AND H. TCHELEPI, *Stochastic formulation for uncertainty analysis of two-phase flow in heterogeneous reservoirs*. *SPE Journal* 5 (2000), pp. 60–70.